

GenCore version 5.1.5

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1090.13 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctgcctgtattttt.....agcgaagctttgaccggat 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum Match 100%

Listing first 45 summaries

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	44.2	356	9	AI475638
2	146.4	18.3	559	14	BM690352
3	145.8	18.2	482	14	R55841
4	145.8	18.2	664	10	BE383336
5	145.8	18.2	676	10	BE260495
6	145.8	18.2	692	10	BE382353

7	145.8	18.2	698	12	BE795434
8	145.8	18.2	726	13	BM009099
9	145.8	18.2	756	13	BE795434
10	145.8	18.2	761	12	BE795434
11	145.8	18.2	774	14	BE795434
12	145.8	18.2	776	14	BE795434
13	145.8	18.2	784	12	BE795434
14	145.8	18.2	791	12	BE795434
15	145.8	18.2	795	14	BE795434
16	145.8	18.2	798	14	BE795434
17	145.8	18.2	799	12	BE795434
18	145.8	18.2	800	12	BE795434
19	145.8	18.2	801	14	BE795434
20	145.8	18.2	802	14	BE795434
21	145.8	18.2	803	14	BE795434
22	145.8	18.2	804	14	BE795434
23	145.8	18.2	805	14	BE795434
24	145.8	18.2	806	14	BE795434
25	145.8	18.2	807	14	BE795434
26	145.8	18.2	808	14	BE795434
27	145.8	18.2	809	14	BE795434
28	145.8	18.2	810	14	BE795434
29	145.8	18.2	811	14	BE795434
30	145.8	18.2	812	14	BE795434
31	145.8	18.2	813	14	BE795434
32	145.8	18.2	814	14	BE795434
33	145.8	18.2	815	14	BE795434
34	145.8	18.2	816	14	BE795434
35	145.8	18.2	817	14	BE795434
36	145.8	18.2	818	14	BE795434
37	145.8	18.2	819	14	BE795434
38	145.8	18.2	820	14	BE795434
39	145.8	18.2	821	14	BE795434
40	145.8	18.2	822	14	BE795434
41	145.8	18.2	823	14	BE795434
42	145.8	18.2	824	14	BE795434
43	145.8	18.2	825	14	BE795434
44	145.8	18.2	826	14	BE795434
45	145.8	18.2	827	14	BE795434

ALIGNMENTS

RESULT 1

AI475638

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI475638 356 bp mRNA linear EST 18-MAR-1999
 tc86d11.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2073045 3',
 mRNA sequence.
 AI475638
 AI475638.1 GI:4328683
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 356)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
 M.D., Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Sequencing by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 2499 Std Error: 0.00
 Seq primer: -40UP from Gibco

1	354	44.2	356	9	AI475638
2	146.4	18.3	559	14	BM690352
3	145.8	18.2	482	14	R55841
4	145.8	18.2	664	10	BE383336
5	145.8	18.2	676	10	BE260495
6	145.8	18.2	692	10	BE382353

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: esc@watson.wustl.edu
Insert Size: 1777
High quality sequence stops: 387 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1777 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 387.

FEATURES

```

location/Qualifiers
1. .482
/organism="Homo sapiens"
/db_xref="GDB:413472"
/db_xref="taxon:9606"
/clon="IMAGE:40931"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: whole brain; Vector: lafmid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
AATCGAAGAATTCCGGCGCGAGGATTTTTTTTTTTT 3'] ;
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."
107 a 139 c 131 g 102 t
3 others
BASE COUNT

```

	Query Match	18.2%	Score 145.8	DB 14	Length 482
	Best Local Similarity	89.1%	Pred. No. 5.6e-20		
	Matches 156	Conservative 1	Mismatches 18	Indels 0	Gaps 0
QY	338	AGTGGCAGTGACTCTCTTCCTCTCTCTCTGACGCCGACACGTCCTCTCTACTGGAC	397		
Db	149	ATTGGAATGTCTAGTGGCACACTTGTCAACATATAGCCCCGACACGTCCTCTCTACTGGAC	208		
QY	398	TGTGGTGAGGGGCACRTTTTGGGCAGCTGTGCCGTCAATTACGGAGACCAAGGTGGACAGGGTC	457		
Db	209	TGTGGTGAGGGGCACATTTGGGCAGCTGTGCCGTCAATTACGGAGACCAAGGTGGACAGGGTC	268		
QY	458	CTGGGCACCCCTGGCTGTGTGTTTGTGTCCCACTCTGACACGAGATCACCCACACGG	512		
Db	269	CTGGGCACCCCTGGCTGTGTGTTTGTGTCCCACTCTGACACGAGATCACCCACACGG	323		

RESULT 4
 BE3833336
 LOCUS
 DEFINITION
 BE3833336
 mRNA
 664 bp linear
 EST 21-JUL-2000
 601298249F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3628308 5',
 mRNA sequence.
 BE3833336
 ACCESSION
 BE3833336.1
 VERSION
 GI:9328701
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 664)
 NTH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 Unpublished (1999)
 JOURNAL
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 COMMENT
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrived by: The I.M.A.G.E. Consortium (LNLN)

DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM313 row: a column: 13
High quality sequence stop: 662.

FEATURES

```

location/Qualifiers
1. .664
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone_id="IMAGE:3628308"
/clone_lib="NIH_MGC.19"
/issue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGACACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."
154 a 199 c 174 g 137 t

```

	Query Match	18.2%;	Score 145.8;	DB 10;	Length 664;
	Best Local Similarity	89.1%;	Pred. No. 5.2e-20;		
	Matches 156; Conservative	1;	Mismatches 18;	Indels 0;	Gaps 0;
Qy	338	AGTGGCAGTGACCTCTCTTCCTCTCTCGAGCCCCGACAGCTCTGTGCTACTGGAC	397		
Dd	394	ATTGGAATGTGAGTGCCACACATTGTCAACATTAAGCCCCGACAGCTCTGTGCTACTGGAC	453		
Qy	398	TGTGGTAGGGGCACRTTTGGCGAGCTGTGCCCGTCATTACGGAGACCAGGTGGACAGGGTC	457		
Dd	454	TGTGGTAGGGGCACATTTGGCGAGCTGTGCCCGTCATTACGGAGACCAGGTGGACAGGGTC	513		
Qy	458	CTGGGCACCTGGCTGCTGTGTTGTGTCCCACTGCCACGCAGATCACCAACGG	512		
Dd	514	CTGGGCACCTGGCTGCTGTGTTGTGTCCCACTGCCACGCAGATCACCAACGG	568		

RESULT 5
BE260495
LOCUS
DEFINITION
601150702f1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:3503184 5',
mRNA linear EST 26-OCT-2000
mRNA sequence.
ACCESSION
BE260495
VERSION
BE260495.1 GI:9131807
KEYWORDS
EST.
SOURCE
human.

REFERENCE	11H-MGC http://mgc.nci.nih.gov/ .
AUTHORS	(bases 1 to 676)
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-rc@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLC1M176 row: d column: 01 High quality sequence stop: 672.

```
source
1. .676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3503184"
/clone_lib="NIH MGC 19"
```


Qy	398	TGTGGTGAAGGCACR TTTGGGACAGCTGTGCCGTCA NTATCGAGACCAGGTGGACAGGGTC	457
Db	68	: :	127
Qy	458	CTGGGCACCCTGGCTGCTGTGTTTGTCGCCACCTGCAGCAGATCACCCACCG	512
Db	128	CTGGGCACCCTGGCTGCTGTGTTTGTCGCCACCTGCAGCAGATCACCCACCG	182
RESULT 8			
BM009099			
LOCUS			
DEFINITION	BM009099	726 bp mRNA linear EST 30-OCT-2001	
ACCESSION	60362945Ef1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5434981 5'		
VERSION	BM009099	mRNA sequence.	
KEYWORDS	BM009099.1 GI:16523453		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 726)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-re@mail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLML at: http://image.llnl.gov Plate: LLCMI912 row: o column: 14 High quality sequence stop: 636. Location/Qualifiers 1..726 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5434981" /clone_lib="NIH_MGC_41" /tissue_type="amelanotic melanoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2: into EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	174 a 195 c 209 g 148 t		
ORIGIN			
Query Match	18.2%; Score 145.8; DB 13; Length 726;		
Best Local Similarity	89.1%; Pred. No. 5.1e-20;		
Matches	156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;		
Qy	338	AGTGCAGTGA CTCTCTTCCTCTCTCTTCCTG CGACCCC CGACAGCTCTCTCTACTGGAC	397
Db	138	ATT CGAAATGT CAGTGCCACACTTGT CAACATAAGCCCC CGACAGCTCTCTCTACTGGAC	197
Qy	398	TGTGGTGAAGGCACR TTTGGGACAGCTGTGCCGTCA NTATCGAGACCAGGTGGACAGGGTC	457
Db	198	: :	257
Qy	458	CTGGGCACCCTGGCTGCTGTGTTTGTCGCCACCTGCAGCAGATCACCCACCG	512
Db	258	CTGGGCACCCTGGCTGCTGTGTTTGTCGCCACCTGCAGCAGATCACCCACCG	312
RESULT 9			

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

NOTE: THIS IS A NIH-PMC LIBRARY.

BASE COUNT	205 a	239 c	248 g	183 t	1 others
------------	-------	-------	-------	-------	----------

```
Query Match      18.2%; Score 145.8; DB 14; Length 876;
Best Local Similarity 89.1%; Pred. No. 4.9e-20;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

Qy 338 AGTGGCAGTGACTCTCTTCCTCTTCTTTCTGACGCCCGACAGTCTCTGCTACTGGAC 397
| | | | | | | | | | | | | | | | | | | | | |
Dβ 230 ATTGGAAATGTCA GTGCCACACTGTCAAACAATAAGCCCGCACAGTCTCTGCTACTGGAC 289
| | | | | | | | | | | | | | | | | | | | | |

Qy 398 TGTGGTGAGGGCACBTTTGGGCAGCTGTGCCGTCTATTACGGAGACCAGGTGGACAGGGTC 457
|||||:|||||
Db 290 TGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCTATTACGGAGACCAGGTGGACAGGGTC 349

Qy	458	CTGGGCA	CCCTGGCTGCTGTTTGTGTCCACCTGCACGACATCACACACGG	512
Dp	350	CTGGGAC	CCCTGGCTGCTGTTTGTGTCCACCTGCACGACATCACACACGG	404

RESULT 13
BG519751
LOCUS BG519751 884 bp mRNA linear EST 02-APR-2001
DEFINITION 602578902F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3503184 5',
mRNA sequence.

ACCESSION	EG519751
VERSION	EG519751.1
KEYWORDS	CI:13515513 EST.

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 884)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D.

The source document: AICC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLNL at: image.llnl.gov
 Plate: LNCM176 row: d column: 01
 High quality sequence stop: 859.

```

FEATURES
source
    high quality sequence stop: 83%.
    Location/Qualifiers
        1..884
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone IMAGE:3503184"
            /clone_lib="NIH MGC_19"
            /tissue type="neuroblastoma"
            /lab_host="DH10B (phage-resistant)"
            /note="organ: brain; Vector: pCIB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGACGAG(G). Library constructed by Ling Hong
            in the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH MGC Library."

```

BASE COUNT	199 a	251 c	249 g	185 t
ORIGIN				

```
Query Match      18.2%; Score 145.8; DB 12; Length 884;
Best Local Similarity 89.1%; Pred. No. 4.9e-20;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

[illegible]

398 TGTGGTGAGGGCACR^{TTT}TGGGCAGCTGTGCCGTCA^{TT}ACGGAGACCAGGTGGACAGGGT

Qy	458	CTGGGCACCCCTGGCTGCTGTTTGTGTCCCACTGCACGAGATCACCACACGG	512
DB	504	TG1GGTGGAGGGGCACATT1GGGCGAGCTGTGCCCCTCAT1ACGAGACACAGGTGGACAGGGT	

DB 564 CTGGGCAACCCCTGGCTGCTGTGTTTGTTGTCCTGCACGAGATCACACACGG 611

RESULT 14

BF311926	BF311926	915 bp	linear	EST 21-NOV-2000
LOCUS	G01897767F1			
DEFINITION	601897767F1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4126864 5', mRNA sequence			

ACCESSION	BF311926	
VERSION	BF311926.1	GI:11259697
KEYWORDS	EST.	
SOURCE	human	

SOURCE ORGANISM

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 915)
TITLE	N1H-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-k@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLC1017 row: f column: 17 High quality sequence stop: 672.

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4126864"
/clone_lib="NIH MGC 19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/NOTE="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

BASE COUNT	206 a	262 c	288 g	159 t
ORIGIN				

Query Match 18.2%; Score 145.8; DB 12; Length 915;
Best Local Similarity 89.1%; Pred. NO. 4.9e-20;
Matches 156: Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 ATGGGCACTGACATCTCTTCTCCTCTTCTTGCGAGCCCCGACACGTCTCTGCTACTGGAC 397

D6 413 ATTCCGAATCTCATCTGCCGCACATTCTCTCACATATAGCCCCGACACCTCTCTGCTACTGGAC 472

Qy	398	TGTGGTGAGGGCA	CRTTTGGGCA	AGCTGTGCGCGT	CAATTACGGAGAC	CCAGGTGCACAGGGTC	457
				:			
Db	473	TGTGGTGAGGGCA	CAATTTGGGCA	AGCTGTGCGCGT	CAATTACGGAGAC	CCAGGTGCACAGGGTC	532

Search completed: May 17, 2003, 18:18:28
Job time : 1096.13 secs

GenCore version 5.1.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 101.551 Seconds
(without alignments)
10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600
Perfect score: 801
Sequence: 1 agtgcctgctctggtatttt.....agcggaagcttggaccggat 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/FCI_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/FCIUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	800.6	100.0	26664	9	US-09-988-626-28
2	800.6	100.0	26664	9	US-09-988-687-28
3	145.8	18.2	2481	9	US-09-988-626-1
4	145.8	18.2	2481	9	US-09-988-687-1
5	145.8	18.2	2892	9	US-09-988-626-225
6	145.8	18.2	2892	9	US-09-988-687-225
7	145.8	18.2	2908	9	US-09-988-626-223
8	145.8	18.2	2908	9	US-09-988-687-223
9	145.8	18.2	2958	9	US-09-988-626-3
10	145.8	18.2	2958	9	US-09-988-687-3
11	138.6	17.3	139	9	US-09-988-626-20
12	138.6	17.3	139	9	US-09-988-687-20
13	102.2	12.8	2470	9	US-09-988-626-221
14	102.2	12.8	2470	9	US-09-988-687-221
15	39.2	4.9	1152	9	US-09-988-626-221
16	39.2	4.9	1152	10	US-09-764-847-1517
17	39.2	4.9	1555	9	US-09-764-847-1517
18	39.2	4.9	1555	9	US-10-092-154-1518
19	38.8	4.8	1732	9	US-09-764-847-1518
					Sequence 28, Appl

20	38.8	4.8	1732	10	US-09-764-853-379	Sequence 379, App
21	38	4.7	10286	9	US-10-239-676-13	Sequence 13, Appl
22	36.8	4.6	671	9	US-10-092-154-296	Sequence 296, App
23	36.8	4.6	671	10	US-09-764-847-296	Sequence 296, App
24	36.8	4.6	1578	10	US-09-822-830A-583	Sequence 583, App
25	36	4.5	1182	10	US-09-922-261-261	Sequence 261, App
26	36	4.5	2454	9	US-10-050-704-93	Sequence 93, Appl
27	36	4.5	2664	9	US-09-809-391-255	Sequence 255, App
28	35.8	4.5	512	9	US-09-918-995-24438	Sequence 24438, A
29	35.8	4.5	2005	10	US-09-887-576-50	Sequence 50, Appl
30	35.8	4.5	8801	9	US-10-239-676-144	Sequence 144, App
31	35.8	4.5	17056	10	US-09-893-238-3	Sequence 3, Appli
32	35.6	4.4	2000	9	US-09-938-842A-3132	Sequence 3132, Ap
33	35.2	4.4	3253	10	US-09-759-359A-1	Sequence 1, Appli
34	35	4.4	499	9	US-10-092-154-1546	Sequence 1546, Ap
35	35	4.4	499	10	US-09-764-847-1546	Sequence 1546, Ap
36	34.8	4.3	416	9	US-09-918-995-3035	Sequence 3035, Ap
37	34.8	4.3	559	9	US-09-918-995-16146	Sequence 16146, A
38	34.4	4.3	478	9	US-09-918-995-16997	Sequence 16997, A
39	34.4	4.3	1119	9	US-10-123-155-352	Sequence 352, App
40	34.4	4.3	513509	9	US-09-754-853A-4	Sequence 4, Appli
41	34.2	4.3	419	10	US-09-960-352-13938	Sequence 13938, A
42	34	4.2	653	9	US-10-184-644-402	Sequence 402, App
43	34	4.2	653	9	US-10-184-644-402	Sequence 402, App
44	34	4.2	802	9	US-10-184-644-312	Sequence 312, App
45	34	4.2	802	9	US-10-184-634-312	Sequence 312, App

ALIGNMENTS

RESULT 1

US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 15883-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature

```
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-626-28

Query Match      100.0%; Score 800.6; DB 9; Length 26664;
Best Local Similarity 100.0%; Pred. No. 6.4e-234;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
Db 21800 AGTGCTGTCTCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy 61 TCATTGTTAGGACTAAATGCTTTAAAGCCTCTTATAAACTTTTTTTTTTTTTTTTGTATGC 120
Db 21860 TCATTGTTAGGACTAAATGCTTTAAAGCCTCTTATAAACTTTTTTTTTTTTTTTTGTATGC 21919

Qy 121 CCAGCCTTTGTGTAGTCTACTTGAAGGGTTTTCAGGGTTCATGATGACTCTTTTGTCTA 180
Db 21920 CCAGCCTTTGTGTAGTCTACTTGAAGGGTTTTCAGGGTTCATGATGACTCTTTTGTCTA 21979

Qy 181 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 240
Db 21980 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 22039

Qy 241 CTCCTCAAAGCAAAGCAGGAGACAGAAATTTCTGCAGTTCCTTCTTGGTCTCTCTCCAA 300
Db 22040 CTCCTCAAAGCAAAGCAGGAGACAGAAATTTCTGCAGTTCCTTCTTGGTCTCTCTCCAA 22099

Qy 301 ACAGACATCAGCTCTGACCATCAGCAGTCTTCTAGTGCAGTACTCTTCTCTCT 360
Db 22100 ACAGACATCAGCTCTGACCATCAGCAGTCTTCTAGTGCAGTACTCTTCTCTCTCT 22159

Qy 361 TCTCTTCTGACGCCGACACGCTCTGCTACTGCTGAGTCTGTGTGAGGGCAGCTTTGGGCA 420
Db 22160 TCTCTTCTGACGCCGACACGCTCTGCTACTGCTGAGTCTGTGTGAGGGCAGCTTTGGGCA 22219

Qy 421 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGCACCCCTGGCTGTGTGT 480
Db 22220 GCTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTCTGGCACCCCTGGCTGTGTGT 22279

Qy 481 TGTGTCCACCTGACGCGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 540
Db 22280 TGTGTCCACCTGACGCGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 22339

Qy 541 ASCCTGGAGGAGGACTGCGACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCTCCGCTT 600
Db 22340 ASCCTGGAGGAGGACTGCGACGTTGAGTTGGCCCTTTGGCTGCGTCTTTTCTCCGCTT 22399

Qy 601 CCAAACTTGCCCGCAGCTTTTGTACTCATCTCTGGCTAGGAAATGGTTTTTGGCAAAAC 660
Db 22400 CCAAACTTGCCCGCAGCTTTTGTACTCATCTCTGGCTAGGAAATGGTTTTTGGCAAAAC 22459

Qy 661 TCAACATAGTCTCTTGGCCGCAAGAAATGCTTCTCTCTCTGTTTCAGTTCCTTTCTCTGC 720
Db 22460 TCAACATAGTCTCTTGGCCGCAAGAAATGCTTCTCTCTCTGTTTCAGTTCCTTTCTCTGC 22519

Qy 721 ACAGGACAGGTTTGTAGTTTACCGACCTTCTCTGAGTCTTGAATCTCACAGGCGCTGCT 780
Db 22520 ACAGGACAGGTTTGTAGTTTACCGACCTTCTCTGAGTCTTGAATCTCACAGGCGCTGCT 22579

Qy 781 CAGCGAAGCTTTGACCGGAT 801
Db 22580 CAGCGAAGCTTTGACCGGAT 22600
```

RESULT 2

```
US-09-988-687-28
; Sequence 28, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-687-28
```

Query Match 100.0%; Score 800.6; DB 9; Length 26664;
Best Local Similarity 100.0%; Pred. No. 6.4e-234;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
Db 21800 AGTGCCTGTCCTGGTATTTTCAACAGAGGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 21859

Qy 61 TCATTGTTAGGACTAAATGCTTTAAAGCCTCTTATAAACTTTTTTTTTTTTTTTTGTATGC 120
Db 21860 TCATTGTTAGGACTAAATGCTTTAAAGCCTCTTATAAACTTTTTTTTTTTTTTTTGTATGC 21919

Qy 121 CCAGCCTTTGTGTAGTCTACTTGAAGGGTTTTCAGGGTTCATGATGACTCTTTTGTCTA 180
Db 21920 CCAGCCTTTGTGTAGTCTACTTGAAGGGTTTTCAGGGTTCATGATGACTCTTTTGTCTA 21979

Qy 181 TAAAGAGGATGACATGTAATCACTTTATGTTTAAATTAATTTGGCTTTTATATTAG 240
```


; Publication No. US20030045704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigan, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
 ; FILE REFERENCE: 2318-258
 ; CURRENT APPLICATION NUMBER: US/09/988,687
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 09/564,805
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 223
 ; LENGTH: 2908
 ; TYPE: DNA
 ; ORGANISM: Pan troglodytes
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2478)
 US-09-988-687-223

Query Match 18.2%; Score 145.8; DB 9; Length 2908;
 Best Local Similarity 89.1%; Pred. No. 7.5e-34;
 Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 AGTGCAGTGACTCTCTTCCTCTCTTCCTGACGCCGACAGCTCTCTGCTACTGGAC 397
 Db 1486 ATTGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGCTCTGCTACTGGAC 1545
 QY 398 TGTGGTAGGGCAGCACTTTGGGAGAGCTGTGCCGTCAATTACGGAGACACAGGTGGACAGGGTC 457
 Db 1546 TGTGGTAGGGCAGCACTTTGGGAGAGCTGTGCCGTCAATTACGGAGACACAGGTGGACAGGGTC 1605
 QY 458 CTGGGCCACCTGGGCTGTGTGTTGTGTCCACCTGCAGCAGAGATCACACACGG 512
 Db 1606 CTGGGCCACCTGGGCTGTGTGTTGTGTCCACCTGCAGCAGAGATCACACACGG 1660

RESULT 9
 US-09-988-626-3
 ; Sequence 3, Application US/09988626
 ; Publication No. US20030044959A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigan, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
 ; FILE REFERENCE: 2318-258
 ; CURRENT APPLICATION NUMBER: US/09/988,626
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 09/564,805
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2958
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 45.0714 Seconds
(without alignments)
5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtccgtctcgtgatttt.....agcgaagcttggaccggat 801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	800.6	100.0	26664	4	US-09-564-805-28
2	145.8	18.2	2481	4	US-09-564-805-1
3	145.8	18.2	2892	4	US-09-564-805-225
4	145.8	18.2	2908	4	US-09-564-805-223
5	145.8	18.2	2958	4	US-09-564-805-3
6	138.6	17.3	139	4	US-09-564-805-20
7	102.2	12.8	2470	4	US-09-564-805-221
8	39.8	5.0	7218	1	US-08-232-463-14
9	39.4	4.9	840	4	US-09-376-728-1
10	36.8	4.6	810	4	US-09-134-001C-1951
11	36.6	4.6	19124	2	US-08-487-826B-13
12	36.6	4.6	4403765	4	US-09-103-840A-2
13	36.6	4.6	4411529	4	US-09-103-840A-1
14	36	4.5	1182	4	US-09-463-697-261
15	36	4.5	2664	4	US-09-149-476-235
16	36	4.5	50000	4	US-09-146-053-4
17	35.8	4.5	17056	4	US-09-245-041-3
18	35	4.4	289	4	US-09-007-005-17
19	35	4.4	289	4	US-09-244-796-17
20	34	4.2	3785	4	US-09-889-718-1
21	34	4.2	7037	4	US-09-853-768-10
22	34	4.2	7037	4	US-09-853-768-3
23	32.8	4.1	6999	1	US-08-276-594A-1
24	32.8	4.1	7056	1	US-08-121-202-1
25	32.8	4.1	7881	2	US-08-751-189-1
26	32.8	4.1	7881	4	US-09-060-836-1
27	32.8	4.1	7881	4	US-09-184-445-1

C 28	32.8	4.1	8241	6	5171844-1	Patent No. 5171844
C 29	32.8	4.1	8967	1	US-08-366-851A-1	Sequence 1, Appli
C 30	32.8	4.1	9009	1	US-07-864-004B-3	Sequence 3, Appli
C 31	32.8	4.1	9009	1	US-08-251-937A-3	Sequence 3, Appli
C 32	32.8	4.1	9009	1	US-08-212-133A-1	Sequence 1, Appli
C 33	32.8	4.1	9009	1	US-08-474-503-1	Sequence 1, Appli
C 34	32.8	4.1	9009	2	US-08-670-707A-1	Sequence 1, Appli
C 35	32.8	4.1	9009	4	US-09-037-601-1	Sequence 1, Appli
C 36	32.8	4.1	9009	4	US-09-315-179-1	Sequence 1, Appli
C 37	32.8	4.1	9009	4	US-09-523-656-1	Sequence 1, Appli
C 38	32.8	4.1	9009	5	PCT-US93-03275-3	Sequence 3, Appli
C 39	32.8	4.1	9009	5	PCT-US94-13200-1	Sequence 1, Appli
C 40	32.4	4.0	630	4	US-08-952-736A-4	Sequence 4, Appli
C 41	32.4	4.0	927	4	US-08-952-736A-1	Sequence 1, Appli
C 42	32.4	4.0	1356	4	US-08-952-736A-2	Sequence 2, Appli
C 43	32.2	4.0	1990	4	US-09-149-476-281	Sequence 281, App
C 44	32.2	4.0	2041	4	US-09-149-476-131	Sequence 131, App
C 45	31.8	4.0	1008	4	US-08-235-836C-130	Sequence 130, App

ALIGNMENTS

RESULT 1

```

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavrigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564.805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5850; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: 8 at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1314, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 2221 and 23879 is A or G.

```

US-09-564-805-28

Query Match 100.0%; Score 800.6; DB 4; Length 26664;
Best Local Similarity 100.0%; Pred. No. 7.6e-232;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGCTGCTGCTGGTATTTTCAACAGAGCTGTGGCCACAGCAATCTGCTAGTGCAGAT 60
DB 21800 AGTGCTGCTGCTGGTATTTTCAACAGAGCTGTGGCCACAGCAATCTGCTAGTGCAGAT 21859

QY 61 TCATTGTTAGGACTAAATGCTTTAAAGCCTCCCTATAAACTTTTTTTTTTTTTTTGTATGC 120
DB 21860 TCATTGTTAGGACTAAATGCTTTAAAGCCTCCCTATAAACTTTTTTTTTTTTTTTGTATGC 21919

QY 121 CCAGCCTTTGTGTAAGTCTACTTGAAGGTTTCAGGGTTCCATGATACCTTTTGCTA 180
DB 21920 CCAGCCTTTGTGTAAGTCTACTTGAAGGTTTCAGGGTTCCATGATACCTTTTGCTA 21979

QY 181 TAAAGAGGATGACATGTAATCAACCTTTATGTTAAATTAATTTGGCTTTTATATTAG 240
DB 21980 TAAAGAGGATGACATGTAATCAACCTTTATGTTAAATTAATTTGGCTTTTATATTAG 22039

QY 241 CTCCTCAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTCTTTGGTCCCTGTCCAA 300
DB 22040 CTCCTCAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTCTTTGGTCCCTGTCCAA 22099

QY 301 AGCAGACATCAGCCTCTGAAACCATGACAGTCTTCTAGTGGCAGTACTCTCTCTCT 360
DB 22100 AGCAGACATCAGCCTCTGAAACCATGACAGTCTTCTAGTGGCAGTACTCTCTCTCT 22159

QY 361 TCTCTTTGACGCCCGACACGCTCTCTGCTACTGAGTCTGTGAGGGGACATTTGGGCA 420
DB 22160 TCTCTTTGACGCCCGACACGCTCTCTGCTACTGAGTCTGTGAGGGGACATTTGGGCA 22219

QY 421 GCTGTGCCGTCTATTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGGCTGCTGTGT 480
DB 22220 GCTGTGCCGTCTATTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGGCTGCTGTGT 22279

QY 481 TGTGTCCACCTGACGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 540
DB 22280 TGTGTCCACCTGACGAGATCACCACAGGTGAGTGTGGGCTGGACCAAAAGCTGG 22339

QY 541 ASCCTGGAGGAGCACTGCCAGCTTGAGTTGCCCTTGGCTGGCTTTTCTCGCTT 600
DB 22340 ASCCTGGAGGAGCACTGCCAGCTTGAGTTGGCCCTTGGCTGGCTTTTCTCGCTT 22399

QY 601 CCAAACTTGGCCAGAGCTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTGCAAAAC 660
DB 22400 CCAAACTTGGCCAGAGCTTTGTTACTCATCTCTGGCTAGGAAATGGTTTTTGCAAAAC 22459

QY 661 TCAACATAGTCTCTTGGCCCAAGAATGCTTCTCTCTCTGTTGAGTTCCTTTCTCTGC 720
DB 22460 TCAACATAGTCTCTTGGCCCAAGAATGCTTCTCTCTCTGTTGAGTTCCTTTCTCTGC 22519

QY 721 ACAGGACAGGTTTGAGTTTACCAGCCTTCTTGAATCTCAGCGGCTGCT 780
DB 22520 ACAGGACAGGTTTGAGTTTACCAGCCTTCTTGAATCTCAGCGGCTGCT 22579

QY 781 CAGCGGAGCTTTGACCGGAT 801
DB 22580 CAGCGGAGCTTTGACCGGAT 22600

RESULT 2

US-09-564-805-1
; Sequence 1, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-1

Query Match 18.2%; Score 145.8; DB 4; Length 2481;
Best Local Similarity 89.1%; Pred. No. 3.8e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 AGTGGCAGTGAATCTCTTCCTCTCTCTCTGCGAGCCCGGACACGCTCTCTCTACTGGAC 397
DB 1486 ATTGAAATGTCAGTGCACACTTGTCAACATAAGCCCGGACACGCTCTCTCTACTGGAC 1545

QY 398 TGTGTGAGGACACRTTTGGCAGCTGTGCCGTCTATTACGAGACCCAGGTGGACAGGTC 457
DB 1546 TGTGTGAGGACACATTTGGCAGCTGTGCCGTCTATTACGAGACCCAGGTGGACAGGTC 1605

QY 458 CTGGCACCCTGGCTGCTGTGTTGTCCTCCACCTGCGACGAGATCACCACGG 512
DB 1606 CTGGCACCCTGGCTGCTGTGTTGTCCTCCACCTGCGACGAGATCACCACGG 1660

RESULT 3

US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-225

Query Match 18.2%; Score 145.8; DB 4; Length 2892;
Best Local Similarity 89.1%; Pred. No. 4.1e-34;
Matches 156; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 338 AGTGGCAGTGAATCTCTTCCTCTCTCTGCGAGCCCGGACACGCTCTCTCTACTGGAC 397
DB 1486 ATTGAAATGTCAGTGCACACTTGTCAACATAAGCCCGGACACGCTCTCTCTACTGGAC 1545

; CURRENT APPLICATION NUMBER: US/09/376,728
; CURRENT FILING DATE: 1999-08-17
; EARLIER APPLICATION NUMBER: US 60/097,242
; EARLIER FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(623)
US-09-376-728-1

Query Match 4.9%; Score 39.4; DB 4; Length 840;
Best Local Similarity 52.8%; Pred. No. 0.034;
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 307 CATCAGCCTCTGAACCATCAGCAGTCTTCTAGTGCAGTACTCTTCTCTCTCTT 366
Db 615 CAACAGCGCGCGCGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 556
QY 367 CTGACGCCCGACACGCTCTCTGCTACTGACCTGTGTGAGGGCACRTTTGGGCGCTG 426
Db 555 CGCGCGGTGGCGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 496
QY 427 CGCTATTACGGACACAGGTGACAGGCTCTGGGCACCC 467
Db 495 CTTGATGACGGCGGCAGCTGTGCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCT 455

RESULT 10
US-09-134-001C-1951/c
; Sequence 1951, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1951
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1951

Query Match 4.6%; Score 36.8; DB 4; Length 810;
Best Local Similarity 69.4%; Pred. No. 0.2;
Matches 50; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 47 CTGCGGTTCAGATTCATTGTTAGGACTAAATGCTTTAAGCCTCTCTATAAAGCTTTT 106
Db 139 CAGCAGTTTCAGATTCATTGATAAGATAAATCCCTAAACATTCCTATAAATGTTTAA 80
QY 107 TTTTCTTTTGTAT 118
Db 79 TTCTTTCTTTAT 68

RESULT 11
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.

; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelisen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13
Query Match 4.6%; Score 36.6; DB 2; Length 19124;
Best Local Similarity 53.1%; Pred. No. 1.4;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 93 ATAAACTTTTTTTTTTTTTTTTTTTTGTATGCCAGCCTTTGTGTAAGTCTACTTGAAGGTT 152
Db 15941 ATAAATTTTTTTTTTTTTTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTT 15882
QY 153 TCAGGGTTCATGGATCTTCTTCTATAAGAGGATGACACATGTAATAATCACCTTTTA 212
Db 15881 AAATTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 15822
QY 213 TGGTTAAATTAATTCGCTTTTATATTA 239
Db 15821 TTATTTAAATAAATTTTTTTTTTTTATTTA 15795
RESULT 12
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          4.6%; Score 36.6; DB 4; Length 4403765;
Best Local Similarity 50.9%; Pred. No. 25;
Matches 84; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 359 CTTCTCTTCTGAGCCCGACACAGTCTCTGTCTACTGGACTGTGTGAGGGACACRTTTGGG 418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435315 CTTGCTTTCGTGGGCTCAATGGGCTCAGCAACTCGACGAGCGCGCGGCTCGGC 435374

QY 419 CAGCTGTGCGGTCATTAGGAGACACAGGTGGACAGGGTCTGGGACACCTGTGCTGCTGTG 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435375 CGCCTCGTCCACACAGGGTCTCGCGGGTTCGCGCGACCGCAGGTCTCTCAACCGCG 435434

QY 479 TTTGTGTCCCACTGCAGCGCAGATCACCACACGGGTGAGTGTGGG 523
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435435 CATGTGCGCCAGAACTGCGCAACCCGCAACCGGACTGGCCGG 435479

RESULT 13
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          4.6%; Score 36.6; DB 4; Length 4411529;
Best Local Similarity 50.9%; Pred. No. 25;
Matches 84; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

QY 359 CTTCTCTTCTGAGCCCGACACAGTCTCTGTCTACTGGACTGTGTGAGGGACACRTTTGGG 418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435233 CTTGCTTTCGTGGGCTCAATGGGCTCAGCAACTCGACGAGCGCGGCGGCTCGGC 435292

QY 419 CAGCTGTGCGGTCATTAGGAGACACAGGTGGACAGGGTCTGGGACACCTGTGCTGCTGTG 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435293 CGCCTCGTCCACACAGGGTCTCGCGGGTTCGCGCGACCGCAGGTCTCTCAACCGCG 435352

QY 479 TTTGTGTCCCACTGCAGCGCAGATCACCACACGGGTGAGTGTGGG 523
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435353 CATGTGCGCCAGAACTGSCCGAACCCGCAACCGGACTGGCCGG 435397

RESULT 14
US-09-461-697-261/c
; Sequence 261, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Putanam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-261

Query Match          4.5%; Score 36; DB 4; Length 1182;
Best Local Similarity 51.9%; Pred. No. 0.44;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 99 TTTTCTTTTCTTTTGTGATGCCAGCCTTTGTGTAAGTCTACTTTGAAAGGTTTCAGGG 158
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1152 TTTTCTTTTCTTTTATTATACCATCAATTTGAAATTTTACAGGAACCAAGATTCAAG 1093

QY 159 TTCCATGATGACTCTTTTGTCTATAAAGAGATGACACATGTAAATCACCCTTTTATGGTTA 218
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1092 CTCCTTAGGTGCTACTGTATTTTATGTTGCACCGACAGACGTCATACACAGTTTCAAT 1033

QY 219 AATTAATTTGGCTTTTATATTAGTCTCTCAAGCAAA 254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1032 AGTAATTTTTCATCATATAGGTTTTCCTTAAAAAAA 997

RESULT 15
US-09-149-476-255/c
; Sequence 255, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
```

/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,662
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,872
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,882
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,637
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,903
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,888
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,879
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,880
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,894
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,911
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,636
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,874
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,910
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,864
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,892
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/057,761
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/047,595
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,599
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,588
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,585
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,594
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,586
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,590
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,594
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,589
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/043,578
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,576
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/047,501
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/043,670
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/056,632
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,664
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,876
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,881
/	EARLIER FILING DATE:	1997-08-22

GenCore version 5.1.5

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 681.842 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664

Perfect score: 501

Sequence: 1 ggtatgggctgtgcgggg.....ttcgccaagttttttgaca 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	469	93.6	1032	13	BM479686
2	303.2	60.5	345	10	AW407520
3	303	60.5	396	9	AA811170
4	302.6	60.4	372	9	AI500718
5	301.4	60.2	478	10	AW592601
6	301.4	60.2	499	10	AW510825

C 7	301.4	60.2	531	10	AW572950
C 8	301.4	60.2	692	14	BQ181829
C 9	301.4	60.2	712	14	BM977567
C 10	299.8	59.8	394	10	AW248468
C 11	299.4	59.8	527	9	AI089646
C 12	298.4	59.6	316	10	AW511765
C 13	298.2	59.5	318	9	AA766184
C 14	298.2	59.5	410	9	AI033108
C 15	297.2	59.3	374	9	AA223338
C 16	296.6	59.2	354	14	BQ030358
C 17	296.6	59.2	390	9	AA605004
C 18	296.6	59.2	406	9	AA583412
C 19	296.6	59.2	416	9	AI141263
C 20	296.6	59.2	433	9	AI357786
C 21	296.6	59.2	461	10	AW592223
C 22	296.6	59.2	479	9	AA716607
C 23	296.6	59.2	489	9	AI803400
C 24	296.6	59.2	489	9	AA534478
C 25	296.6	59.2	531	9	AI937465
C 26	296.6	59.2	536	14	BM681558
C 27	296.6	59.2	614	13	BI870047
C 28	296.2	59.1	491	9	AA676661
C 29	295.6	59.0	477	9	AI804749
C 30	294.6	58.8	409	12	BF477438
C 31	294.6	58.8	439	9	AI201492
C 32	294.6	58.8	494	12	BE858252
C 33	294.6	58.8	535	9	AI158921
C 34	294.6	58.8	536	9	AI153280
C 35	293.6	58.6	372	14	D29176
C 36	293.6	58.6	481	9	AI145928
C 37	293.6	58.6	541	9	AI148489
C 38	293.4	58.6	306	9	AA862956
C 39	291.8	58.2	409	9	AI831155
C 40	291.6	58.2	352	14	R44712
C 41	290.2	57.9	488	13	BM353522
C 42	289.6	57.8	329	14	R58845
C 43	289.6	57.8	398	9	AA948352
C 44	288.6	57.6	414	9	AA134386
C 45	287.2	57.3	367	14	BQ549604

ALIGNMENTS

RESULT 1
BM479686
LOCUS BM479686 1032 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6418543 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502530
5', mRNA sequence.
ACCESSION BM479686
VERSION BM479686.1 GI:18528728
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1032)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12140 Row: n column: 03
High quality sequence stop: 720.
Location/Qualifiers
1. .1032

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5502530"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
253 a 234 c 295 g 247 t 3 others
ORIGIN
Query Match 93.6%; Score 469; DB 13; Length 1032;
Best Local Similarity 97.6%; Pred. No. 2.8e-131;
Matches 489; Conservative 0; Mismatches 5; Indels 7; Gaps 1;
QY 1 GGTATGGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
Db |||||
QY 170 GGTATGGAGCTGTGCGGAGGCTTGGGTTCCCATAAAGCACTAGTCTATAGATGCTCTT 229
Db |||||
QY 61 AGGACTGTGCTGTCACAGCGCGGGCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db |||||
QY 230 AGGACTGTGCTGTCACAGCTGCGGGCAGAGGCTGCCACACGGAAGCAAGCAGATGA 289
Db |||||
QY 121 ACTAATTTCAATTTCAAGCAGTTTTTAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db |||||
QY 290 ACTAATTTCAATTTCAAGCAGTTTTTAAGAGTCTTGGAAACAGACGCGGCACCTTTC 349
Db |||||
QY 181 CTCTAATCAGCAAAAGTGATTCCTTCGACACAGAGACAGCAGAGTAAACAGATCAGTG 240
Db |||||
QY 350 CTCTAATCAGCAAAATGATTCCTTCGACACAGAGACAGCAGAGTAAACAGATCAGTG 409
Db |||||
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTTGAGTTTG 300
Db |||||
QY 410 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTTGAGTTTG 469
Db |||||
QY 301 CAATTTGTAGTCTTTTGTCTCTCTGCTGCTACAGACAGAGGCTGCTGTGTCACC 360
Db |||||
QY 470 CAATTTGTAGTCTTTTGTCTCTCTGCTGCTACAGACAGAGGCTGCTGTGTCACC 522
Db |||||
QY 361 ACCTTTGGAGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db |||||
QY 523 ACCTTTGGAGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 582
Db |||||
QY 421 AAGAAAGTCAAGTCCCGTGTAGTGAACCTCTGGAACCTGCTCTCAGAGACCCCTT 480
Db |||||
QY 583 AAGAAAGTCAAGTCCCGTGTAGTGAACCTCTGGAACCTGCTCTCAGAGACCCCTT 642
Db |||||
QY 481 ATTCGCCAAGTCTTTTGACA 501
Db |||||
QY 643 ATTCGCCAAGTCTTTTGACA 663
Db |||||

RESULT 2
AW407520/c
LOCUS AW407520 345 bp mRNA linear EST 16-FEB-2000
DEFINITION UI-HF-BMO-adl-b-07-0-UI.r1 NIH_MGC_38 Homo sapiens cDNA clone
IMAGE:3061957 5', mRNA sequence.
ACCESSION AW407520
VERSION AW407520.1 GI:6926577
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 345)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
Location/Qualifiers
1. 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3061957"
/clone_lib="NIH_MGC_38"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/note="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(2.5-3.5Kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
73 a 94 c 82 g 96 t
BASE COUNT
ORIGIN

Query Match 60.5%; Score 303.2; DB 10; Length 345;
Best Local Similarity 99.0%; Pred. No. 3.7e-81;
Matches 305; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTATGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
Db |||||
QY 308 GGTATGAGCTGTGCGGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 249
Db |||||
QY 61 AGGACTGTGCTGGCACAGCGCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db |||||
QY 248 AGGACTGTGCTGGCACAGCGCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 189
Db |||||
QY 121 ACTAATTTCAATTTCAAGGCGAGTTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db |||||
QY 188 ACTAATTTCAATTTCAAGGCGAGTTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 129
Db |||||
QY 181 CTCTAATCCAGCAAGTGATTCCCTGCACACCCAGAGCAAGCAGAGTAAACAGGATCAGTG 240
Db |||||
QY 128 CTCTAATCCAGCAAGTGATTCCCTGCACACCCAGAGCAAGCAGAGTAAACAGGATCAGTG 69
Db |||||
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTTGAGTTTG 300
Db |||||
QY 68 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTTGAGTTTG 9
Db |||||
QY 301 CAATTTGTG 308
Db |||||
QY 8 CCCTCGTG 1

RESULT 3
AA811170/c
LOCUS AA811170 396 bp mRNA linear EST 19-FEB-1998
DEFINITION ob42c03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1334020 3',
mRNA sequence.
ACCESSION AA811170
VERSION AA811170.1 GI:2880781
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 396)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

ACCESSION AW592601
 VERSION AW592601.1 GI:7279786
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 478)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 451.
 FEATURES
 source
 1..478
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2934808"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 92 a 132 c 122 g 132 t
 ORIGIN

Query Match 60.2%; Score 301.4; DB 10; Length 478;
 Best Local Similarity 99.7%; Pred. No. 1.6e-80;
 Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GGTATGGAGCTGTGGCGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
 Db 305 GGTATGGAGCTGTGGCGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 246
 Qy 61 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACAGCAAGCAGATGA 120
 Db 245 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACAGCAAGCAGATGA 186
 Qy 121 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTGGAAACAGACGCGGCACCTTTC 180
 Db 185 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTGGAAACAGACGCGGCACCTTTC 126
 Qy 181 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAAACAGATCAGTG 240
 Db 125 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAAACAGATCAGTG 66
 Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
 Db 65 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 6
 Qy 301 CAA 303
 Db 5 CAA 3

RESULT 6
 AW510825/c
 LOCUS hd40b11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 DEFINITION 499 bp mRNA linear EST 03-MAR-2000

IMAGE:2911965 3', mRNA sequence.
 ACCESSION AW510825
 VERSION AW510825.1 GI:7148903
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 499)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 470.
 FEATURES
 source
 1..499
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2911965"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 94 a 142 c 126 g 137 t
 ORIGIN

Query Match 60.2%; Score 301.4; DB 10; Length 499;
 Best Local Similarity 99.7%; Pred. No. 1.7e-80;
 Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GGTATGGAGCTGTGGCGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
 Db 303 GGTATGGAGCTGTGGCGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 244
 Qy 61 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACAGCAAGCAGATGA 120
 Db 243 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACAGCAAGCAGATGA 184
 Qy 121 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTGGAAACAGACGCGGCACCTTTC 180
 Db 183 ACTAATTTTCATTTCAAGGCAGTCTTTTAAAGAAAGTCTTGGAAACAGACGCGGCACCTTTC 124
 Qy 181 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAAACAGATCAGTG 240
 Db 123 CTCCTAATCCAGCAAAAGTATTCCTTCGACACAGAGCAAGCAGAGTAAACAGATCAGTG 64
 Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 300
 Db 63 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG 4
 Qy 301 CAA 303
 Db 3 CAA 1

RESULT 7
 AW572950/c
 LOCUS AW572950
 DEFINITION 531 bp mRNA linear EST 13-MAR-2000


```

DEFINITION      hf17h05.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE: 2932185 3', mRNA sequence.
ACCESSION       AW572950
VERSION         AW572950.1  GI:7237683
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 531)
AUTHORS         NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT         Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  This clone is available royalty-free through LNL; contact the
                  IMAGE Consortium (info@image.llnl.gov) for further information.
                  Seq primer: -40UP from Gibco
                  High quality sequence stop: 459.
FEATURES        Location/Qualifiers
                 1..531
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:2932185"
                 /clone_lib="Soares_NFL_T_GBC_S1"
                 /lab_host="DH10B"
                 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                 Equal amounts of plasmid DNA from three normalized
                 libraries (fetal lung NBHL19W, testis NHR, and B-cell
                 NCI-CGAP GCB1) were mixed, and ss circles were made in
                 vitro. Following HAP purification, this DNA was used as
                 tracer in a subtractive hybridization reaction. The driver
                 was PCR-amplified cDNAs from pools of 5,000 clones made
                 from the same 3 libraries. The pools consisted of
                 I.M.A.G.E. clones 297480-302087, 682632-687239,
                 726408-728711, and 729096-731399. Subtraction by Bento
                 Soares and M. Fatima Bonaldo. "
BASE COUNT      99 a 154 c 136 g 142 t
ORIGIN
Query Match     60.2%; Score 301.4; DB 10; Length 531;
Best Local Similarity 99.7%; Pred. No. 1.8e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 60
Db 305 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 246

QY 61 AGGACTGGTGGCCCTGGCACAGCGCGCGGCGGAGGCTGCCACACGGAAGCAAGCATGA 120
Db 245 AGGACTGGTGGCCCTGGCACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCATGA 186

QY 121 ACTAATTTTCATTTCAAGGAGTCTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db 185 ACTAATTTTCATTTCAAGGAGTCTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 126

QY 181 CTCCTAATCCAGCAAGTGTATCCCTGCACACGAGAGGAGTACAGGATCAGTG 240
Db 125 CTCCTAATCCAGCAAGTGTATCCCTGCACACGAGAGGAGTACAGGATCAGTG 66

QY 241 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGTTTG 300
Db 65 GGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTCAGCTGCAATAAAGATTGTTTG 6

QY 301 CAA 303
Db 5 CAA 3

RESULT 8
BQ181829/c

```

```

LOCUS           BQ181829
DEFINITION      UI-H-EUO-azv-m-08-0-UI.s1 NCI_CGAP_Carl Homo sapiens cDNA clone
IMAGE: 5854255 3', mRNA sequence.
ACCESSION       BQ181829
VERSION         BQ181829.1  GI:20357321
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 692)
AUTHORS         NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT         Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue Procurement: Dr. Jose Mercuende
                  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                  Clone Distribution: Clone distribution information can be found
                  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                  Seq primer: M13 FORWARD
                  POLYA=yes.
FEATURES        Location/Qualifiers
                 1..692
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:5854255"
                 /clone_lib="NCI_CGAP_Carl"
                 /tissue_type="Osteoarthritic Cartilage"
                 /dev_stage="Adult"
                 /lab_host="DH10B (Life Technologies)"
                 /note="Organ: Knee; Vector: pT7T3-Pac (Pharmacia) with a
                 modified polylinker; Site 1: EcoR I; Site 2: Not I;
                 NCI_CGAP_Carl is a cDNA library containing the following
                 tissue(s): Osteoarthritic Cartilage The library was
                 constructed according to Bonaldo, Lennon and Soares,
                 Genome Research, 6:791-806, 1996. First strand cDNA
                 synthesis was primed with an oligo-dT primer containing a
                 Not I site. Double stranded cDNA was ligated to an EcoR I
                 adaptor, digested with Not I, and cloned directionally
                 into pT7T3-Pac vector. The oligonucleotide used to prime
                 the synthesis of first-strand cDNA contains a library tag
                 sequence that is located between the Not I site and the
                 (dT)18 tail. The sequence tag for this library is
                 TGATCAGCT.
                 TAG_LIB=UI-H-EUO
                 TAG_TISSUE=osteoarthritic cartilage
                 TAG_SEQ=TGATCAGCT"
BASE COUNT      127 a 204 c 174 g 187 t
ORIGIN
Query Match     60.2%; Score 301.4; DB 14; Length 692;
Best Local Similarity 99.7%; Pred. No. 2.1e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 60
Db 319 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCCATAGCACTAGTCTATAGATGCTCTT 260

QY 61 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCATGA 120
Db 259 AGGACTGGTGGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAAGCAAGCATGA 200

QY 121 ACTAATTTTCATTTCAAGGAGTCTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 180
Db 199 ACTAATTTTCATTTCAAGGAGTCTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTC 140

QY 181 CTCCTAATCCAGCAAGTGTATCCCTGCACACGAGAGGAGTACAGGATCAGTG 240
Db 139 CTCCTAATCCAGCAAGTGTATCCCTGCACACGAGAGGAGTACAGGATCAGTG 80

```

```

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 300
|||||
Db 79 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTCAATAAAGATTGAGTTTG 20

Qy 301 CAA 303
|||
Db 19 CAA 17

RESULT 9
LOCUS BM977567/c
DEFINITION UI-CF-EN1-aef-c-19-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
          712 bp mRNA linear EST 21-MAR-2002
ACCESSION BM977567
VERSION UI-CF-EN1-aef-c-19-0-UI 3', mRNA sequence.
KEYWORDS EST.
SOURCE BM977567.1 GI:19596113
        human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
        discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: McCray, PB
          University of Iowa
          2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
          Tel: 319 356 4866
          Fax: 319 356 7171
          Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             source
    Location/Qualifiers
        1..712
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="UI-CF-EN1-aef-c-19-0-UI"
            /clone_lib="UI-CF-EN1"
            /tissue_type="Primary Lung Cystic Fibrosis Epithelial
            Cells"
            /dev_stages="Adult"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
            modified polylinker; Site 1: EcoR I; Site 2: Not I;
            UI-CF-EN1 is a normalized cDNA library containing the
            following tissue(s): Primary Lung Cystic Fibrosis
            Epithelial Cells. The library was constructed according to
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. First strand cDNA synthesis was primed with an
            oligo-dT primer containing a Not I site. Double stranded
            cDNA was ligated to an EcoR I adaptor, digested with Not
            I, and cloned directionally into pT7T3-Pac vector. The
            oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tag for this library is CTGCTCAGGT.
            TAG LIB=UI-CF-EN1
            TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
            6hr to LPS 24h
            TAG_SEQ=CTGCTCAGGT"

BASE COUNT          130 a      211 c      176 g      194 t      1 others
ORIGIN

```

```

Query Match          60.2%   Score 301.4;   DB 14;   Length 712;
Best Local Similarity 99.7%   Pred. No. 2.2e-80;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
|||||
Db 319 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 260

Qy 61 AGGACTGGTGGCTGGCACAGCCGCGGGCGAGGCTGCCACACGGGAAGCAAGCAGATGA 120
|||||
Db 259 AGGACTGGTGGCTGGCACAGCCGCGGGCGAGGCTGCCACACGGGAAGCAAGCAGATGA 200

Qy 121 ACTAATTTCAATTTCAAGCGAGTGTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTTC 180
|||||
Db 199 ACTAATTTCAATTTCAAGCGAGTGTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTTC 140

Qy 181 CTCTAATCCAGCAAAAGTGATTCCCTCGCACACACAGAGACAAGCAGAGTAACAGGATCAGTG 240
|||||
Db 139 CTCTAATCCAGCAAAAGTGATTCCCTCGCACACACAGAGACAAGCAGAGTAACAGGATCAGTG 80

Qy 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 300
|||||
Db 79 GGTCTAAGTGTCCGAGACTTAAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 20

Qy 301 CAA 303
|||
Db 19 CAA 17

RESULT 10
LOCUS AW248468/c
DEFINITION 2820640.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820640 3',
          mRNA sequence.
ACCESSION AW248468.
VERSION AW248468.1 GI:6591461
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: 2820640.5prime
          Contact: Robert Strausberg, Ph.D.
          Email: cgabs-x@mail.nih.gov
          Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
          Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
          Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
          project Clone distribution: MGC clone distribution information can
          be found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
          Scores: PHRED from University of Washington Genome Center
          Trimming: cross_match from University of Washington Genome Center
          PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
          Drosophila Genome Project. University of Washington Genome Center:
          http://www.genome.washington.edu/Polyadenylation/ Based upon the
          presence of a XhoI site followed by a run of 14 or more T residues
          at the beginning of the sequence, this cDNA insert was
          polyadenylated.
          Plate: LLCMA row: L column: 17
          High quality sequence stop: 213.
          Location/Qualifiers
              1..394
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:2820640"
                  /clone_lib="NIH_MGC_7"
                  /tissue_type="small cell carcinoma"
                  /cell_line="MGC3"

```

/lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 80 a 103 c 97 g 114 t
 ORIGIN

Query Match 59.8%; Score 299.8; DB 10; Length 394;
 Best Local Similarity 99.3%; Pred. No. 4.4e-80;
 Matches 301; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGTATGGAGCTGTGCCAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
 DB 311 GGTATGGAGCTGTGCCAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 252
 QY 61 AGGACTGTGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAACAGCAGATGA 120
 DB 251 AGGACTGTGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAACAGCAGATGA 192
 QY 121 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAAGTCTTTGAAACAGACGCGGACCTTTC 180
 DB 191 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAAGTCTTTGAAACAGACGCGGACCTTTC 132
 QY 181 CTCCTAATCCAGCAAGTGTTCCTCCACACAGACAGCAGAGTAAACAGATCAGTG 240
 DB 131 CTCCTAATCCAGCAAGTGTTCCTCCACACAGACAGCAGAGTAAACAGATCAGTG 72
 QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
 DB 71 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 12
 QY 301 CAA 303
 DB 11 CAA 9

RESULT 11
 AI089646/c
 LOCUS AI089646 527 bp mRNA linear EST 18-AUG-1998
 DEFINITION qbi607.x1 Soares pregnant uterus_NbHPU Homo sapiens cDNA clone IMAGE:1696476 3', mRNA sequence.
 ACCESSION AI089646
 VERSION AI089646
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 527)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 463.
 Location/Qualifiers
 1..527
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1696476"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"

FEATURES
 source

/note="Organ: uterus; Vector: pT7T3-Pac; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGGAAGAAATTCGCGCGCCCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."
 BASE COUNT 99 a 153 c 136 g 138 t 1 others
 ORIGIN

Query Match 59.8%; Score 299.4; DB 9; Length 527;
 Best Local Similarity 99.7%; Pred. No. 7.1e-80;
 Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGTATGGAGCTGTGCCAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
 DB 301 GGTATGGAGCTGTGCCAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 242
 QY 61 AGGACTGTGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAACAGCAGATGA 120
 DB 241 AGGACTGTGCTGGCACAGCGCGGCGGAGGCTGCCACACGGAACAGCAGATGA 182
 QY 121 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAAGTCTTTGAAACAGACGCGGACCTTTC 180
 DB 181 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAAGTCTTTGAAACAGACGCGGACCTTTC 122
 QY 181 CTCCTAATCCAGCAAGTGTTCCTCCACACAGACAGCAGAGTAAACAGATCAGTG 240
 DB 131 CTCCTAATCCAGCAAGTGTTCCTCCACACAGACAGCAGAGTAAACAGATCAGTG 62
 QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
 DB 61 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2
 QY 301 C 301
 DB 1 C 1

RESULT 12
 AWS11765/c
 LOCUS AWS11765 316 bp mRNA linear EST 03-MAR-2000
 DEFINITION xu76f03.x1 NCI_CGAP_Kid8 Homo sapiens cDNA clone IMAGE:2807645 3', mRNA sequence.
 ACCESSION AWS11765
 VERSION AWS11765.1 GI:7149843
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 316)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/iresources.shtml
 Seq primer: -40UP from Gibco
 High quality sequence stop: 239.
 Location/Qualifiers
 1..316
 /organism="Homo sapiens"

FEATURES
 source

```

/db_xref="taxon:9606"
/clone="IMAGE:2807645"
/clone_lib="NCI_CGAP_Kid8"
/tissue_type="renal cell tumor"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.2 kb. Life Technologies catalog #:
11524-014"
BASE COUNT      69 a   84 c   72 g   90 t   1 others
ORIGIN

Query Match      59.6%; Score 298.4; DB 10; Length 316;
Best Local Similarity 99.3%; Pred. No. 1e-79;
Matches 299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTCGAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 60
Db 301 GGTATGGAGCTGTCGAGGCTTGGGCTCCACATNAGCAGTCTATAGATGCTCTT 242

Qy 61 AGGACTGTGCTGCGACAGCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 120
Db 241 AGGACTGTGCTGCGACAGCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 182

Qy 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 181 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 122

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGCAAGCAGAGTAAAGGATCAGTG 240
Db 121 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGCAAGCAGAGTAAAGGATCAGTG 62

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 300
Db 61 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 2

Qy 301 C 301
Db 1 C 1

RESULT 13
AA766184/c
LOCUS      318 bp      mRNA      linear      EST 08-FEB-1998
DEFINITION oal2f09.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:130477 3',
            mRNA sequence.
ACCESSION  AA766184
VERSION     AA766184.1 GI:2817422
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 318)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1277 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 288.
            Location/Qualifiers

source
1. .318
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1304777"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-,
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      71 a   82 c   70 g   95 t
ORIGIN

Query Match      59.5%; Score 298.2; DB 9; Length 318;
Best Local Similarity 99.0%; Pred. No. 1.2e-79;
Matches 300; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 60
Db 310 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAAGCAGTCTATAGATGCTCTT 251

Qy 61 AGGACTGTGCTGCGACAGCGCGGCGGAGGCTGCCACAGGAAGCAGATGA 120
Db 250 AGGACTGTGCTGCGACAGCGCGGCGGAGGCTGCCACAGGAAGCAGATGA 191

Qy 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 190 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 131

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGCAAGCAGAGTAAAGGATCAGTG 240
Db 130 CTCTAATCCAGCAAGTGTATCCCTGCACACAGAGCAAGCAGAGTAAAGGATCAGTG 71

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 300
Db 70 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTACGCTGCAATAAGATTGATTG 11

Qy 301 CAA 303
Db 10 CAA 8

RESULT 14
AI033108/c
LOCUS      410 bp      mRNA      linear      EST 28-AUG-1998
DEFINITION ow98g08.sl Soares fetal liver spleen INFUS_S1 Homo sapiens cDNA
            clone IMAGE:1654910 3', mRNA sequence.
ACCESSION  AI033108
VERSION     AI033108.1 GI:3254061
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 410)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES

```

Insert Length: 773 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 368.

FEATURES

Location/Qualifiers
1..410
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1654910"
/clone_lib="Soares_fetal_liver_spleen_INFLS_s1"
/sex="male"
/dev_stages="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AATGGAGAAATTAATTAAGATCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 88 a 110 c 101 g 111 t
ORIGIN
Query Match 59.5%; Score 298.2; DB 9; Length 410;
Best Local Similarity 99.0%; Pred. No. 1.4e-79;
Matches 300; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGGCGAGGCTTGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
DB 305 GGTATGGAGCTGTGGCGAGGCTTGGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 246
QY 61 AGGACTGGTCTGTCACACCGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 120
DB 245 AGGACTGGTCTGTCACACCGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 186
QY 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 180
DB 185 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 126
QY 181 CTCTAATCCAGCAAGTGTATTCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 240
DB 125 CTCTAATCCAGCAAGTGTATTCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 66
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 300
DB 65 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 6
QY 301 CAA 303
DB 5 CAA 3

RESULT 15
AA223338/c

LOCUS AA2233338 374 bp mRNA linear EST 12-MAR-1998
DEFINITION zr05h05.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
CDNA clone IMAGE:650649 3', mRNA sequence.
ACCESSION AA223338
VERSION AA223338.1 GI:1843862
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 374)
AUTHORS Hillier, L., Allen, M., Bowles, T., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project

TITLE

JOURNAL
COMMENT

Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 3011 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

Location/Qualifiers
1..374
/organism="Homo sapiens"
/db_xref="GDB:5277238"
/db_xref="taxon:9606"
/clone="IMAGE:650649"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stages="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb.
Uni-ZAP XR vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTT 3' "

BASE COUNT 78 a 98 c 85 g 112 t 1 others
ORIGIN
Query Match 59.3%; Score 297.2; DB 9; Length 374;
Best Local Similarity 98.7%; Pred. No. 2.6e-79;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGGCGAGGCTTGGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 60
DB 312 GGTATGGAGCTGTGGCGAGGCTTGGGCTCCCATAAAGCACTAGTCTATAGATGCTCTT 253
QY 61 AGGACTGGTCTGTCACACGCGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 120
DB 252 AGGACTGGTCTGTCACACGAGCTGCGGCGAGGCTGCCACACGAGCAAGCAGATGA 193
QY 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 180
DB 192 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGAAACAGACGCGGCACCTTTC 133
QY 181 CTCTAATCCAGCAAGTGTATTCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 240
DB 132 CTCTAATCCAGCAAGTGTATTCCTGCACACGAGACAAAGCAGATGAACAGATCAGTG 73
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 300
DB 72 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCACTGCAATTAAGATTGAGTTTG 13
QY 301 CAA 303
DB 12 CAA 10

Search completed: May 17, 2003, 18:18:22
Job time : 689.842 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 63.5169 Seconds
(without alignments)
10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664
Perfect score: 501
Sequence: 1 ggtatgagctgtccgagg.....ctcccaagttttttgaca 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	26664	9	US-09-988-626-28
2	501	100.0	26664	9	US-09-988-687-28
3	303	60.5	655	9	US-09-988-626-27
4	303	60.5	655	9	US-09-988-687-27
5	303	60.5	2958	9	US-09-988-626-3
6	303	60.5	2958	9	US-09-988-687-3
7	296.6	59.2	2908	9	US-09-988-626-223
8	296.6	59.2	2908	9	US-09-988-687-223
9	254.6	50.8	2892	9	US-09-988-626-225
10	254.6	50.8	2892	9	US-09-988-687-225
11	33.6	6.7	480	10	US-09-864-761-15427
12	32.2	6.4	513509	9	US-09-754-853A-4
13	32	6.4	1089	10	US-09-974-300-2320
14	32	6.4	3779	10	US-09-880-107-2145
15	31.4	6.3	2049	10	US-09-887-576-655
16	31.2	6.2	1101	9	US-10-102-806-206
17	31.2	6.2	2327	9	US-09-803-391-107
18	30.8	6.2	529	10	US-09-917-800A-1269
19	30.6	6.1	350	10	US-09-960-352-8591

c	20	30.6	6.1	367	10	US-09-960-352-5310	Sequence 5310, Ap
c	21	30.6	6.1	406	10	US-09-960-352-9329	Sequence 9329, Ap
	22	30.6	6.1	421	9	US-10-091-504-415	Sequence 415, App
	23	30.6	6.1	421	10	US-09-764-869-415	Sequence 415, App
c	24	30.6	6.1	463	10	US-09-960-352-1733	Sequence 1733, Ap
c	25	30.6	6.1	38844	12	US-10-060-333-3	Sequence 3, Appli
	26	30.2	6.0	294	10	US-09-294-093B-2539	Sequence 2539, Ap
	27	30.2	6.0	728	9	US-10-079-854-54	Sequence 54, Appl
	28	30.2	6.0	728	10	US-09-764-878-54	Sequence 114, App
	29	30.2	6.0	1595	9	US-10-108-605-114	Sequence 452, App
c	30	30	6.0	293	10	US-09-294-093B-452	Sequence 452, App
	31	30	6.0	418	9	US-09-918-995-2942	Sequence 2942, Ap
c	32	30	6.0	515	10	US-09-815-343-704	Sequence 704, App
c	33	30	6.0	779	10	US-09-910-943-602	Sequence 602, App
c	34	30	6.0	1901	10	US-09-834-975-952	Sequence 952, App
	35	30	6.0	4355	12	US-10-044-090-10	Sequence 10, Appl
	36	30	6.0	8586	10	US-09-960-253-170	Sequence 170, App
	37	29.8	5.9	576	10	US-09-924-035A-806	Sequence 806, App
	38	29.8	5.9	180557	12	US-10-003-806-6	Sequence 6, Appli
	39	29.8	5.9	180557	12	US-10-003-806-9	Sequence 9, Appli
c	40	29.6	5.9	392	9	US-09-764-891-6473	Sequence 6473, Ap
c	41	29.6	5.9	392	9	US-09-764-891-6473	Sequence 6474, Ap
c	42	29.6	5.9	544	10	US-09-815-242-20	Sequence 20, Appl
c	43	29.4	5.9	555	9	US-10-123-155-72	Sequence 72, Appl
c	44	29.4	5.9	564	10	US-09-864-761-7967	Sequence 7967, Ap
c	45	29.4	5.9	577	10	US-09-867-701-2000	Sequence 2000, Ap

ALIGNMENTS

RESULT 1

US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5583-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature

Db 26644 ATTCGCAAGTCTTTTGACA 26664
|||||

RESULT 3

US-09-988-626-27
; Sequence 27, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA_signal
; LOCATION: (636)..(641)
US-09-988-626-27

Query Match 60.5%; Score 303; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 353 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 412
Qy 61 AGGACTGGTGCCTGGCAGACGCGCGCGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 413 AGGACTGGTGCCTGGCAGACGCGCGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 473 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 240
Db 533 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 241 GGTCTAAGTGTCCGAGACTTAAAGAAAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 473 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 240
Db 533 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 241 GGTCTAAGTGTCCGAGACTTAAAGAAAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 593 GGTCTAAGTGTCCGAGACTTAAAGAAAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 301 CAA 303
Db 653 CAA 655

RESULT 4

US-09-988-687-27
; Sequence 27, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA_signal
; LOCATION: (636)..(641)
US-09-988-687-27

Query Match 60.5%; Score 303; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
Db 353 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 412
Qy 61 AGGACTGGTGCCTGGCAGACGCGCGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 413 AGGACTGGTGCCTGGCAGACGCGCGCGGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 473 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 181 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 240
Db 533 CTCCTAATCCAGCAAAAGTATTCCTCCACACGAGAGAGGCTGCCACACGGAAGCAAGCAGATGA 472
Qy 241 GGTCTAAGTGTCCGAGACTTAAAGAAAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 180
Db 593 GGTCTAAGTGTCCGAGACTTAAAGAAAGTCTTTTAAAGAGTCTTTGGAACAGACGCGGCACCTTTTC 532
Qy 301 CAA 303
Db 653 CAA 655

RESULT 5

US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626

; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match 60.5%; Score 303; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 4.3e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db |||||
Qy 61 AGGACTGTGCTGCCACAGCGCGGCGGAGGCTGCCACAGCAAGCAAGCATGA 120
Db |||||
Qy 121 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAGTCTTTGGAACACAGCGGCGCACCTTTC 180
Db |||||
Qy 181 CTCTAATCCAGCAAGTATTCCTCTGCACACAGAGAGTAAAGAGTCTTTGGAACACAGCGGCGCACCTTTC 240
Db |||||
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 300
Db |||||
Qy 301 CAA 303
Db |||||
Qy 2956 CAA 2958

RESULT 6
US-09-988-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-687-3

Query Match 60.5%; Score 303; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 4.3e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db |||||
Qy 61 AGGACTGTGCTGCCACAGCGCGGCGGAGGCTGCCACAGCAAGCAAGCATGA 120
Db |||||
Qy 121 ACTAATTTTCATTTCAAGCAGTCTTTAAAGAGTCTTTGGAACACAGCGGCGCACCTTTC 180
Db |||||
Qy 181 CTCTAATCCAGCAAGTATTCCTCTGCACACAGAGAGTAAAGAGTCTTTGGAACACAGCGGCGCACCTTTC 240
Db |||||
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGAGTTTG 300
Db |||||
Qy 301 CAA 303
Db |||||
Qy 2956 CAA 2958

RESULT 7
US-09-988-626-223
; Sequence 223, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-223

Query Match 59.2%; Score 296.6; DB 9; Length 2908;
Best Local Similarity 98.7%; Pred. No. 6.7e-91;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	GGTATGGAGCTGTGCCGAGGCTTGGCTCCCACTAATAGCACTAGTCTATAGATGCCCTCTT	60
Db	2606	GGTATGGAGCTGTGCCAAGGCTTGGCTCCCACTAATAGCACTAGTCTATAGATGCCCTCTT	2665
Qy	61	AGGACTGTGTGCTTGGCACAGCCGCGGGCCAGAGGCTGCCACAGGAAGCAAGCAGATGA	120
Db	2666	AGGACTGGTGTGCTTGGCACAGCCGCGGACAGGAGCTGCCACAGGAAGCAAGCAGATGA	2725
Qy	121	ACTAATTTTCATTTCAAGGCAGATTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTC	180
Db	2726	ACTAATTTTCATTTCAAGGCAGATTTTAAAGAGGCTTGGAAACAGACGGCAGCACCTTTC	2785
Qy	181	CTCTAATCCAGCAAGTGATTCCTTGCAACACAGAGACAAGCAGAGTAAACAGGATCAAGT	240
Db	2786	CTCTAATCCAGCAAGTGATTCCTTGCAACACAGAGACAAGCAGAGTAAACAGGATCAAGT	2845
Qy	241	GGTCTAAGTGTCGGAGACCTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG	300
Db	2846	GGTCTAAGTGTCGGAGACCTTAACGAAATAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG	2905
Qy	301	CAA	303
Db	2906	CAA	2908

RESULT 8

```

US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223

```

Qy	181	CTCTAATCCAGCAAGATGATTCCTCGACACACAGAGACAACAGAGATTAACAGGATCAAGT	240
Db	2786	CTCTAATCCAGCAAGATGATTCCTCGACACAGAGACAACAGAGATTAACAGGATCAAGT	2845
Qy	241	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATTAAGATTTGAGTTTG	300
Db	2846	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATTAAGATTTGAGTTTG	2905
Qy	301	CAA 303	
Db	2906	CAA 2908	

RESULT 9
 US-09-988-626-225
 ; Sequence 225, Application US/09988626
 ; Publication No. US20030044959A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
 ; FILE REFERENCE: 2318-258
 ; CURRENT APPLICATION NUMBER: US/09/988,626
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 09/564,805
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 225
 ; LENGTH: 2892
 ; TYPE: DNA
 ; ORGANISM: Gorilla gorilla
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2478)
 US-09-988-626-225

Query Match	50.8%	Score 254.6;	DB 9;	Length 2892;
Best Local Similarity	93.4%;	Pred. No. 1.7e-76;		
Matches 283;	Conservative	0;	Mismatches 4;	Indels 16;
				Gaps
Qy	1	GGTATGAGCTGTGCGAGGCTTGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT	60	
Db	2606	GGTATGAGCTGTGCGAGGCTTAGGCTCCACATAAGCACTAGTCTATATA-----	2655	
Qy	61	AGGACTGTGTCCTTGGCAGACCGCGGGCCAGGAGCTGCCACACGGAAGCAAGCAGATGA	120	
Db	2656	-----GGTGCCTTGGCAGACCGCGGGCAGGAGCTGCCACACGGAAGCAAGCAGATGA	2709	
Qy	121	ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGGAAACAGACGGCGGCACCTTTC	180	
Db	2710	ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTTGGAAACAGACGGCGGCACCTTTC	2769	
Qy	181	CTCTAATCCAGCAAAATGATTCCTTGCAACACAGAGACAAGCAGAGTAAACAGGATCAGTG	240	
Db	2770	CTCTAATCCAGCAAAATGATTCCTTGCAACACAGAGACAAGCAGAGTAAACAGGATCAGTG	2829	
Qy	241	GGTCTAAGTGTCCGAGACACTTAAACGAAAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG	300	
Db	2830	GGCTAAGTGTCCGAGACTTAAACGAAAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTG	2889	
Qy	301	CAA 303		
Db	2890	CAA 2892		

```
RESULT 10
US-09-988-687-225
; Sequence 225, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-225

Query Match      50.8%; Score 254.6; DB 9; Length 2892;
Best Local Similarity 93.4%; Pred. No. 1.7e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;

Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 2606 GGTATGGAGCTGTGCGAGGCTTAGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60

Qy 61 AGGACTGTGCTGCGACAGCCGGCGGCGAGGCTGCCACAGCAAGCAAGCAGATGA 120
Db 2656 -----GGTGCTGTGCGACAGCCGGCGGCGAGGCTGCCACAGCAAGCAGATGA 2709

Qy 121 ACTAATTTCAATTTAAGGCAGTTTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTC 180
Db 2710 ACTAATTTCAATTTAAGGCAGTTTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTC 2769

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGCAGACAGCAGATTAACAGGATCAGTG 240
Db 2770 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGCAGACAGCAGATTAACAGGATCAGTG 2829

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTCTCAATAAGATTGAGTTTG 300
Db 2830 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTCTCAATAAGATTGAGTTTG 2889

Qy 301 CAA 303
Db 2890 CAA 2892

RESULT 11
US-09-864-761-15427/c
; Sequence 15427, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15427
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009230.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
US-09-864-761-15427

Query Match      6.7%; Score 33.6; DB 10; Length 480;
Best Local Similarity 47.2%; Pred. No. 0.44;
Matches 102; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 176 CTTTCCTCTAATCCAGCAAGTGTATCCCTGCGACACAGCAGACAGCAGATTAACAGAT 235
Db 344 CTTGAAGCTATTTGGAAATGGTGTAGCTTTTAAAGAAAAAGCAAGTGAACACAGTAA 285

Qy 236 CAGTGGGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTCAATAAAGATTGA 295
Db 284 ACAGAGACAAAAAATTTTGGAAATCAACATGGTACAAATGACCACTTTTAACAGATGA 225

Qy 296 GTTTGCAATTTGTAGATTCTTTTGGCTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGT 355
Db 224 AACAAACTTATAGCTTTTCATCCTTACAATTACTGTTTCTAGAGAGCTAGATTACTCA 165

Qy 356 GCACACCTTGGAGAAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
Db 164 TCAGGCTTTTGGGAATTTTCTTTTATTTCCCTTTTGT 129
```

```

RESULT 12
US-09-754-853A-4
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: Soybean Cyst Nematode Resistance
; CURRENT APPLICATION NUMBER: US/09/754,853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111805)..(113968),(114684)..(115204)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

```

	Query Match	6.4%	Score 32.2;	DB 9;	Length 513509;
	Best Local Similarity	66.7%;	Pred. No. 69;		
	Matches 46;	Conservative 0;	Mismatches 23;	Indels 0;	Gaps 0;
QY	3	TATGGAGCTGTGCGAGGCTTGGGCTCCACATAAAGCACTAGTCATATAGATGCCCTCTTAG	62		
Db	503895	TTTGAAGTTGTGCAAAAGGATGTGCTCACAAAGGCCCTTAGTACATTGAGTGCTGTTAG	503954		
QY	63	GACTGGTGTC	71		
Db	503955	TACTGCTGC	503963		

```

RESULT 13
US-09-974-300-2320/c
; Sequence 2320, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2320
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2320

```

Query Match	6.4%	Score 32;	DB 10;	Length 1089;
Best Local Similarity	56.7%	Pred. NO. 2.5;		
Matches 59;	Conservative	0;	Mismatches 45;	Indels 0;
Gaps	0;			

QY 270 AGTATTTGAGCTGCAATAAAGATTGAGTTTGCAATTGTGAGTTCCTTCTCCTGC 329

```

Db      527 ATTCTCCCATCTATAAACACATGTGTCATAAAATTTTTCAGCTTTTGCTCTTCGGA 468
QY      330 TGCTGTCACAGACAGCGGGTCTGCTGTCGCACCACCTTGGAGAAG 373
Db      467 CGCTCTCCTCACGAGNACCTTTTTTCTTCTCGAATAAAAAG 424

RESULT 14
US-09-880-107-2145
; Sequence 2145, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2145
; LENGTH: 3779
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L07077
US-09-880-107-2145

```

	Query Match	6.4%	Score 32	DB 10	Length 3779
	Best Local Similarity	54.2%	Pred. No. 5		
	Matches 65	Conservative 0	Mismatches 55	Indels 0	Gaps 0
Qy	132	TTCAAGGCGAGTTTTTAAGAAAGCTCTTGGAACAGACGCGCGCACCTTTCCCTCTCTAATCCAG	191		
Db	1422	TTGTAGGCAACTGTTTTGGATTTCTGGGGAATCGAATGTTGAATCCTTTACTACAACTCAGG	1481		
Qy	192	CAAAAGTGATTCCTTCGCACACCAGAGACAAGCAGAGAGTAACAGGATCAGTGGGTCTAAAGTGT	251		
Db	1482	CATATTTCTTGTTAGAGAAGCGAGCAAAACAGAGGAGGTAGATCAGGTCTGGGAAGAGT	1541		

```

RESULT 15
US-09-887-576-655/c
; Sequence 655, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 655

```

```
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-655

Query Match      6.3%; Score 31.4; DB 10; Length 2049;
Best Local Similarity 56.2%; Pred. No. 5.7;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy  274 TTTCAGCTGCAATAAAGATTGAGTTTGCAATTGTGAGTTCCTTTTGGCTTCCTCCTCCTGCT 333
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  116 TTTATGTCGTAAGGACCATTTGGCTTTTGAAAAAATCAGATCTCTGAATGCAGGTGGTGCA 57

Qy  334 GCTACAGACGAGGCTCTGCTGTGCACCACTTGGAGAGGCTCTC 378
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   56 GCGACAGCACAGAGCTTATGACCACAAACACAGAAAGATTGCTCGC 12
```

Search completed: May 18, 2003, 00:39:58
Job time : 384.517 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 28.1908 Seconds
(without alignments)
5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664

Perfect score: 501

Sequence: 1 ggtatggagctgtgcgagg.....ttcgcaagctgttttgaca 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PTUS COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	501	100.0	26564	4	US-09-564-805-28
2	303	60.5	655	4	US-09-564-805-27
3	303	60.5	2958	4	US-09-564-805-3
4	296.6	59.2	2908	4	US-09-564-805-223
5	254.6	50.8	2892	4	US-09-564-805-225
6	33.2	6.6	6042	1	US-08-261-822A-1
7	33.2	6.6	6042	5	PCT-US95-07744A-1
8	33.2	6.6	6172	2	US-08-819-288-1
9	33.2	6.6	6172	4	US-09-400-348-1
10	31.2	6.2	2327	4	US-09-149-476-107
11	31.2	6.2	8916	4	US-09-579-181-11
12	31.2	6.2	9354	4	US-09-579-181-10
13	30.6	6.1	787	4	US-08-943-731-200
14	30.6	6.1	20084	4	US-08-943-731-5
15	30.6	6.1	38844	4	US-09-734-675-3
16	30.2	6.0	35828	4	US-09-449-218D-17
17	30	6.0	1834	1	US-08-592-126-90
18	29.8	5.9	7218	1	US-08-232-463-14
19	29.6	5.9	630	4	US-08-180-371-17
20	29.6	5.9	1677	2	US-08-684-101-1
21	29.6	5.9	1677	4	US-09-205-814-1
22	29.4	5.9	1342	3	US-08-961-083-181
23	29.4	5.9	3048	1	US-08-188-228-47
24	29.4	5.9	3048	1	US-08-332-643-41
25	29.4	5.9	3048	1	US-08-332-638-47
26	29.4	5.9	4453	2	US-08-843-530B-17
27	29.4	5.9	6867	4	US-08-961-527-192

c	28	29	5.8	506	1	US-08-469-802B-7	Sequence 7, Appli
c	30	29	5.8	506	2	US-08-267-803B-7	Sequence 7, Appli
c	31	29	5.8	49795	4	US-09-453-702B-60	Sequence 60, Appli
c	32	28.8	5.7	847	4	US-09-142-565-5	Sequence 5, Appli
c	33	28.8	5.7	3186	1	US-08-761-258-8	Sequence 8, Appli
c	34	28.8	5.7	3186	2	US-08-977-306-8	Sequence 8, Appli
c	35	28.8	5.7	5053	2	US-08-685-576-2	Sequence 2, Appli
c	36	28.6	5.7	3395	4	US-09-103-478-3	Sequence 3, Appli
c	37	28.6	5.7	7560	4	US-09-103-931C-3	Sequence 4, Appli
c	38	28.6	5.7	7560	4	US-09-193-931C-4	Sequence 4, Appli
c	39	28.6	5.7	11531	1	US-08-068-945A-1	Sequence 1, Appli
c	40	28.6	5.7	11531	1	US-08-442-806-1	Sequence 1, Appli
c	41	28.4	5.7	877	2	US-08-394-152A-44	Sequence 44, Appli
c	42	28.2	5.6	391	4	US-09-641-638-419	Sequence 419, App
c	43	28.2	5.6	2718	4	US-09-651-656-14	Sequence 14, Appli
c	44	28.2	5.6	2718	4	US-09-650-855-14	Sequence 14, Appli
c	45	28	5.6	603	4	US-09-134-001C-1226	Sequence 1226, Ap

ALIGNMENTS

RESULT 1

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5850; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1314, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 2221 and 23879 is A or G.


```
Db 2716 AGGACTGGTCCCTGGCACACGCCGGCCAGAGGCTGCCACACGGAACGACAGATGA 2775
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTGAAACAGACGGCGGCACCTTTC 180
Db 2776 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTGAAACAGACGGCGGCACCTTTC 2835
QY 181 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 240
Db 2836 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 2895
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2896 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2955
QY 301 CAA 303
Db 2956 CAA 2958

RESULT 4
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-223

Query Match 59.2%; Score 296.6; DB 4; Length 2908;
Best Local Similarity 98.7%; Pred. No. 7.6e-91;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGAGCTGTGCCAAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 2665
QY 61 AGGACTGGTCCCTGGCACACGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTCCCTGGCACACGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 2725
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTTAAAGAACTCTTTTAAAGAACTCTTTT 180
Db 2726 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTTAAAGAACTCTTTTAAAGAACTCTTTT 2785
QY 181 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 240
Db 2786 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 2845
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2905
```

```
QY 301 CAA 303
Db 2906 CAA 2908

RESULT 5
US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-225

Query Match 50.8%; Score 254.6; DB 4; Length 2892;
Best Local Similarity 93.4%; Pred. No. 1.6e-76;
Matches 283; Conservative 0; Mismatches 4; Indels 16; Gaps 1;

QY 1 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 60
Db 2606 GGTATGAGCTGTGCCAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTT 2655
QY 61 AGGACTGGTCCCTGGCACACGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2656 -----GGTGCCTGGCACACGCCGGCCAGAGGCTGCCACACGGAAGCAAGCAGATGA 2709
QY 121 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTTAAAGAACTCTTTTAAAGAACTCTTTT 180
Db 2710 ACTAATTTTCATTTCAAGGAGTCTTTTAAAGAACTCTTTTAAAGAACTCTTTTAAAGAACTCTTTT 2769
QY 181 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 240
Db 2770 CTCCTAATCCAGCAAGTATTCCTCGCACACACGAGCAACGAGAGATTAACAGGATCAGTG 2829
QY 241 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 300
Db 2830 GGTCTAAGTGTCCGAGACTTAAACGAAATAAGTATTTTTCAGCTGCAATAAAGATTGAGTTTG 2889
QY 301 CAA 303
Db 2890 CAA 2892

RESULT 6
US-08-261-822A-1
; Sequence 1, Application US/08261822A
; Patent No. 5650553
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261.822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-261-822A-1
;
; Query Match
; Best Local Similarity 6.6%; Score 33.2; DB 1; Length 6042;
; Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; Qy 200 TTCCTGTCACACAGACAGACAGAGTAAACAGGATCAGTGGTCTAAGTGTCCGAGACT 259
; Db 618 TTCCTGGAAGATCTGAATCGGTAGATCATACGGGATCTTGGCAATTTGCTTTTCGT 677
;
; Qy 260 TAACGAAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCATTTGCAATTTGCTTTTGC 319
; Db 678 CAGCGTTACGATCTTTAGCTTCAGTTAGTTGAAATTTGATTTTTCAGCTTATC 737
;
; Qy 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
; Db 738 TTCCTTTTGTGCTGCTTCTACTAAGATC 767
;
; RESULT 7
; PCT-US95-07744A-1
; Sequence 1, Application PC/TUS9507744A
; GENERAL INFORMATION:
; APPLICANT: Trustees of The University of Pennsylvania
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07744A
; FILING DATE: 15-JUNE-1995
;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261.822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; PCT-US95-07744A-1
;
; Query Match
; Best Local Similarity 51.3%; Score 33.2; DB 5; Length 6042;
; Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; Qy 200 TTCCTGTCACACAGACAGACAGAGTAAACAGGATCAGTGGTCTAAGTGTCCGAGACT 259
; Db 618 TTCCTGGAAGATCTGAATCGGTAGATCATACGGGATCTTGGCAATTTGCTTTTCGT 677
;
; Qy 260 TAACGAAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCATTTGCAATTTGCTTTTGC 319
; Db 678 CAGCGTTACGATCTTTAGCTTCAGTTAGTTGAAATTTGATTTTTCAGCTTATC 737
;
; Qy 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
; Db 738 TTCCTTTTGTGCTGCTTCTACTAAGATC 767
;
; RESULT 8
; US-08-819-288-1
; Sequence 1, Application US/08819288
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,288
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: UPN-2949
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,822
; FILING DATE: June 17, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; PCT-US95-07744A-1
;
; Query Match
; Best Local Similarity 51.3%; Score 33.2; DB 5; Length 6042;
; Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
;
; Qy 200 TTCCTGTCACACAGACAGACAGAGTAAACAGGATCAGTGGTCTAAGTGTCCGAGACT 259
; Db 618 TTCCTGGAAGATCTGAATCGGTAGATCATACGGGATCTTGGCAATTTGCTTTTCGT 677
;
; Qy 260 TAACGAAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCATTTGCAATTTGCTTTTGC 319
; Db 678 CAGCGTTACGATCTTTAGCTTCAGTTAGTTGAAATTTGATTTTTCAGCTTATC 737
;
; Qy 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
; Db 738 TTCCTTTTGTGCTGCTTCTACTAAGATC 767
;
; RESULT 8
; US-08-819-288-1
; Sequence 1, Application US/08819288
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,288
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: UPN-2949
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 6172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-819-288-1

Query Match
Best Local Similarity 51.3%; Pred. No. 0.68;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 200 TTCCTGACACAGACAGACAGAGTACAGGATCAGTGGGTCTAAAGTGTCCGAGACT 259
DB 750 TTCCTTGAAGATCTGAATGCGTAGATCATACGGGATCTTTGCAATTTTGTGCTTTTCGT 809
QY 260 TAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCATTTGAGTTCTTTTTCG 319
DB 810 CAGCGTACAGATCTTTAGCTTCAGTTAGTTGAAATTTGTAATTTTGTGAGCTTATC 869
QY 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
DB 870 TTCCTTTGCTGCTTTCATACCTAGATC 899

RESULT 9
US-09-400-348-1
; Sequence 1, Application US/09400348
; Patent No. 6355778
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6355778ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/400,348
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,288
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: UPN-2949
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-400-348-1

Query Match
Best Local Similarity 6.6%; Score 33.2; DB 4; Length 6172;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 200 TTCCTGACACAGACAGACAGTAAACAGGATCAGTGGGTCTAAAGTGTCCGAGACT 259
DB 750 TTCCTTGAAGATCTGAATGCGTAGATCATACGGGATCTTTGCAATTTTGTGCTTTTCGT 809
QY 260 TAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTGCATTTGAGTTCTTTTTCG 319
DB 810 CAGCGTACAGATCTTTAGCTTCAGTTAGTTGAAATTTGTAATTTTGTGAGCTTATC 869
QY 320 TTCCTCTGCTGCTGTACAGACAGGGTC 349
DB 870 TTCCTTTGCTGCTTTCATACCTAGATC 899

RESULT 10
US-09-149-476-107
; Sequence 107, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: *60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,613	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,582	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,596	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,612	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,632	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,601	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/043,580	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,568	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,311	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,671	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,674	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,669	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,312	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/048,974	EARLIER FILING DATE: 1997-06-06	EARLIER APPLICATION NUMBER: 60/056,886	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,877	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,889	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,893	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,630	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,878	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,662	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,872	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,882	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,637	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,903	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,888	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,879	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,880	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,894	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,911	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,636	EARLIER FILING DATE: 1997-08-22
----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------	----------------------------------------	---------------------------------

1	EARLIER	FILING DATE:	1997-08-22
2	EARLIER	APPLICATION NUMBER:	60/056,874
3	EARLIER	FILING DATE:	1997-08-22
4	EARLIER	APPLICATION NUMBER:	60/056,910
5	EARLIER	FILING DATE:	1997-08-22
6	EARLIER	APPLICATION NUMBER:	60/056,864
7	EARLIER	FILING DATE:	1997-08-22
8	EARLIER	APPLICATION NUMBER:	60/056,631
9	EARLIER	FILING DATE:	1997-08-22
10	EARLIER	APPLICATION NUMBER:	60/056,845
11	EARLIER	FILING DATE:	1997-08-22
12	EARLIER	APPLICATION NUMBER:	60/056,892
13	EARLIER	FILING DATE:	1997-08-22
14	EARLIER	APPLICATION NUMBER:	60/057,761
15	EARLIER	FILING DATE:	1997-08-22
16	EARLIER	APPLICATION NUMBER:	60/047,595
17	EARLIER	FILING DATE:	1997-05-23
18	EARLIER	APPLICATION NUMBER:	60/047,599
19	EARLIER	FILING DATE:	1997-05-23
20	EARLIER	APPLICATION NUMBER:	60/047,588
21	EARLIER	FILING DATE:	1997-05-23
22	EARLIER	APPLICATION NUMBER:	60/047,585
23	EARLIER	FILING DATE:	1997-05-23
24	EARLIER	APPLICATION NUMBER:	60/047,586
25	EARLIER	FILING DATE:	1997-05-23
26	EARLIER	APPLICATION NUMBER:	60/047,590
27	EARLIER	FILING DATE:	1997-05-23
28	EARLIER	APPLICATION NUMBER:	60/047,594
29	EARLIER	FILING DATE:	1997-05-23
30	EARLIER	APPLICATION NUMBER:	60/047,589
31	EARLIER	FILING DATE:	1997-05-23
32	EARLIER	APPLICATION NUMBER:	60/047,593
33	EARLIER	FILING DATE:	1997-05-23
34	EARLIER	APPLICATION NUMBER:	60/047,614
35	EARLIER	FILING DATE:	1997-05-23
36	EARLIER	APPLICATION NUMBER:	60/043,578
37	EARLIER	FILING DATE:	1997-04-11
38	EARLIER	APPLICATION NUMBER:	60/043,576
39	EARLIER	FILING DATE:	1997-04-11
40	EARLIER	APPLICATION NUMBER:	60/047,501
41	EARLIER	FILING DATE:	1997-05-23
42	EARLIER	APPLICATION NUMBER:	60/043,670
43	EARLIER	FILING DATE:	1997-04-11
44	EARLIER	APPLICATION NUMBER:	60/056,632
45	EARLIER	FILING DATE:	1997-08-22
46	EARLIER	APPLICATION NUMBER:	60/056,664
47	EARLIER	FILING DATE:	1997-08-22
48	EARLIER	APPLICATION NUMBER:	60/056,876
49	EARLIER	FILING DATE:	1997-08-22
50	EARLIER	APPLICATION NUMBER:	60/056,881
51	EARLIER	FILING DATE:	1997-08-22
52	EARLIER	APPLICATION NUMBER:	60/056,909
53	EARLIER	FILING DATE:	1997-08-22
54	EARLIER	APPLICATION NUMBER:	60/056,875
55	EARLIER	FILING DATE:	1997-08-22
56	EARLIER	APPLICATION NUMBER:	60/056,862
57	EARLIER	FILING DATE:	1997-08-22
58	EARLIER	APPLICATION NUMBER:	60/056,887
59	EARLIER	FILING DATE:	1997-08-22
60	EARLIER	APPLICATION NUMBER:	60/056,908
61	EARLIER	FILING DATE:	1997-08-22
62	EARLIER	APPLICATION NUMBER:	60/048,964
63	EARLIER	FILING DATE:	1997-06-06
64	EARLIER	APPLICATION NUMBER:	60/057,650
65	EARLIER	FILING DATE:	1997-09-05
66	EARLIER	APPLICATION NUMBER:	60/056,884
67	EARLIER	FILING DATE:	1997-08-22
68	EARLIER	APPLICATION NUMBER:	60/057,669
69	EARLIER	FILING DATE:	1997-09-05
70	EARLIER	APPLICATION NUMBER:	60/049,610
71	EARLIER	FILING DATE:	1997-06-13
72	EARLIER	APPLICATION NUMBER:	60/061,060
73	EARLIER	FILING DATE:	1997-10-00

Query Match 6.2%; Score 31.2; DB 4; Length 2327;
Best Local Similarity 35.9%; Pred. No. 1.8;
Matches 33; Conservative 28; Mismatches 31; Indels 0; Gaps 0;
QY 327 TGTGCTGTACAGACAGGCTGTGTCACACCTTGGAGAGGCTCTGTGCTCT 386
Db 1576 WSCTKSWGCTCWCCKSRSTGRMKGMRTCTAGAAATRGYRGAACMYIKSGCTKQWGG 1635
QY 387 AGTGTGGCAGCTGCTGTGTTACCGGGTGGCTT 418
Db 1636 AAKKSGGASGAGGACGACCTGCATTGCTT 1667

RESULT 11

US-09-579-181-11/c
; Sequence 11, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chrivia, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 8916
; TYPE: DNA
; ORGANISM: Human
US-09-579-181-11

Query Match 6.2%; Score 31.2; DB 4; Length 8916;
Best Local Similarity 50.7%; Pred. No. 4;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 13 TGGCAGGCTTGGCTCCACATAAGCACTAGTCTATAGATGCTCTTAGGACTGGTGCC 72
Db 6900 TGCCAATGCTGGGCACACAGCTCTGCTCAGGCCCTCAGCCCAAGGGGAGATTAGTAC 6841
QY 73 TGGCAGCGCGGGCCAGAGGCTCCACACGGAAGCAGATGAACCTAATTTCAAT 132
Db 6840 TGAGGCAGAGATGGGCACAGAGGTGGACCAAGCAGAGAGGAGGAGTTACAAG 6781
QY 133 TCAAGCAGTCTTTAAAGAGCTTTGA 160
Db 6780 ACAGGTTTGGGCTGGAGCGGTGTATGA 6753

RESULT 12

US-09-579-181-10/c
; Sequence 10, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chrivia, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 9354
; TYPE: DNA
; ORGANISM: Human
US-09-579-181-10

Query Match 6.2%; Score 31.2; DB 4; Length 9354;
Best Local Similarity 50.7%; Pred. No. 4.2;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 13 TGGCAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTTAGGACTGGTGCC 72
Db 7338 TGCCAATGCTGGGCACACAGCTCTGCTCAGGCCCTCAGCCCAAGGGGAGATTAGTAC 7279
QY 73 TGGCAGCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACCTAATTTCAAT 132
Db 7278 TGAGGCAGAGATGGGCACAGAAAGGTGGACCAAGCAAGAGAGAGGAGGAGTTACAAG 7219
QY 133 TCAAGCAGTCTTTAAAGAGCTTTGA 160
Db 7218 ACAGGTTTGGGCTGGAGCGGTGTATGA 7191

RESULT 13

US-08-943-731-200/c
; Sequence 200, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY PH.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-943-731-200

Query Match 6.1%; Score 30.6; DB 4; Length 787;
Best Local Similarity 52.8%; Pred. No. 1.6;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 377 TCTGTGCTGTAGTGTGGCAGCTGCTGGTACCCGGTGGCTTGGAGAAGTCAGCTCCCG 436
|||||
DB 340 TCTGGCTGAGGTGGGACCCAGCAGAGAGGAGTGGCCGACGGTGCCCGGACCC 281
|||||

QY 437 TGTAGTGAGCACCTCTGGAACTCTCTCAGAGAGCACCTTATTCGCCAAGTCCTTTT 496
|||||
DB 280 TGGGAGTGCCCACTCTGCCCAGTGGTGCTACCCACCCCTTACTTCCCGTGCTCT 221
|||||

QY 497 TGACA 501
|||
DB 220 GTGCA 216
|||

RESULT 14

US-08-943-731-5/c
; Sequence 5, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943.731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY Ph.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-5

Query Match 6.1%; Score 30.6; DB 4; Length 20084;
Best Local Similarity 52.8%; Pred. No. 10;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 377 TCTGTGCTGTAGTGTGGCAGCTGCTGGTACCCGGTGGCTTGGAGAAGTCAGCTCCCG 436
|||||
DB 14308 TCTGGCTGAGGTGGGACCCAGCAGAGAGGAGTGGCCGACGGTGCCCGGACCC 14249
|||||

QY 437 TGTAGTGAGCACCTCTGGAACTCTCTCAGAGAGCACCTTATTCGCCAAGTCCTTTT 496
|||||
DB 14248 TGGAGTGCCCACTCTGCCAGTGGTGCTACCCACCCCTTACTTCCCGTGCTCT 14189
|||||

QY 497 TGACA 501
|||
DB 14188 GTGCA 14184
|||

RESULT 15

US-09-734-675-3/c
; Sequence 3, Application US/09734675
; Patent No. 6365391
; GENERAL INFORMATION:
; APPLICANT: WESTER, MARION et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000862
; CURRENT APPLICATION NUMBER: US/09/734,675
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 38844
; TYPE: DNA
; ORGANISM: Human
US-09-734-675-3

Query Match 6.1%; Score 30.6; DB 4; Length 38844;
Best Local Similarity 46.5%; Pred. No. 15;
Matches 99; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 113 GCAGATCAACTAATTTCAATTTCAAGGCAGTTTTTAAAGAAGTCTTTGAAAACAGACGCGG 172
|||||
DB 31579 GTATAAGATAAAACAAAACCAATGGGACATCTGGAATAAATCTAGAGACCTTAAACCG 31520
|||||

QY 173 CACCTTTCTCTAATCCAGCAAGTGATTCCTGCACACAGAGAGAGAGTACAG 232
|||||

DB 31519 ATTCCTGCCTATGTGAGTGAAAGTTATTCCTAGTTCCCTTAATGGTAGACTGAGAATCAA 31460
|||||

QY 233 GATCAGTGGGTCTAAGTGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGAT 292
|||||

DB 31459 AAGCTTATCTTTATATTTCTTGATGTAATAAATGTTATAGAAATCTCTGCAATAGAT 31400
|||||

QY 293 TGATTTGCAATGTGAGTCTTTTGTCTTCTC 325
|||||

DB 31399 TTAGTCATCTCAGAAAAGCCCTTTTGCAATTTTC 31367
|||||

Search completed: May 17, 2003, 16:47:42
Job time : 72.1908 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1298.06 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_26164_26664
Perfect score: 501
Sequence: 1 99tatggagctgtgcgagg.....ttcgccaagtctttttgaca 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_man.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	118788	9	AC005277	AC005277 Homo sapi
2	303	60.5	2908	9	AF304370	AF304370 Homo sapi
3	301.4	60.2	2163	9	AK074244	AK074244 Homo sapi
4	301.4	60.2	2997	9	BC001939	BC001939 Homo sapi
5	301.4	60.2	3006	9	BC004158	BC004158 Homo sapi
6	298.4	59.8	2907	9	AK094687	AK094687 Homo sapi
7	296.6	59.2	2908	9	AF308698	AF308698 Pan trogl
8	296.6	59.2	2992	6	AX405825	AX405825 Sequence
9	294.6	58.8	2734	9	AK094333	AK094333 Homo sapi
10	294.6	58.8	2793	9	AK094012	AK094012 Homo sapi
11	294.6	58.8	2976	9	AK001392	AK001392 Homo sapi
C 12	286.4	57.2	429	11	G60268	G60268 SHGC-130964
13	254.6	50.8	2893	9	AF308694	AF308694 Gorilla g
14	158	31.5	386	6	AX069570	AX069570 Sequence
C 15	133.4	26.6	137	11	G42927	G42927 WIAF-176-ST
C 16	95.4	19.0	97	6	AX326639	AX326639 Sequence
C 17	95.4	19.0	97	6	AX326672	AX326672 Sequence
C 18	93.8	18.7	97	6	AX326638	AX326638 Sequence
19	93.8	18.7	97	6	AX326659	AX326659 Sequence
C 20	48.2	9.6	186836	2	AC097911	AC097911 Rattus no
21	42.2	8.4	34593	10	AF348157	AF348157 Mus muscu
22	42.2	8.4	196566	10	AL663045	AL663045 Mouse DNA
23	41.4	8.3	125020	9	AF429315	AF429315 Homo sapi
24	40.2	8.0	16402	4	MIBMCG	AF429315 Homo sapi
C 25	39.8	7.9	141115	2	AL450991	AL450991 Homo sapi
26	39.8	7.9	154008	30	AC058799	AC058799 Homo sapi
27	39.8	7.9	187127	9	AC097359	AC097359 Homo sapi
C 28	37.8	7.5	125020	9	AF429315	AF429315 Homo sapi
29	37	7.4	338116	2	AC087159	AC087159 Mus muscu
C 30	36.8	7.3	159524	9	CNS01DUO	AL133305 Human chr
31	36.8	7.3	160276	2	AC013529	AC013529 Homo sapi
C 32	36.6	7.3	146638	9	AC010347	AC010347 Homo sapi
C 33	36.6	7.3	177902	9	AC026992	AC026992 Homo sapi
34	36.4	7.3	83675	9	AC004830	AC004830 Homo sapi
C 35	36.4	7.3	218542	2	AL731655	AL731655 Mus muscu
C 36	35.6	7.1	220989	2	AC105707	AC105707 Rattus no
C 37	35.6	7.1	252576	2	AC111963	AC111963 Rattus no
C 38	35.4	7.1	105256	2	AC126180	AC126180 Rattus no
39	35.4	7.1	166386	2	AC126082	AC126082 Rattus no
C 40	35.4	7.1	212335	2	AC126941	AC126941 Mus muscu
41	35.2	7.0	55965	9	AC104653	AC104653 Homo sapi
42	35.2	7.0	125685	2	AC069268	AC069268 Homo sapi
43	35.2	7.0	145465	2	AC069399	AC069399 Homo sapi
44	35.2	7.0	163496	9	AC092567	AC092567 Homo sapi
C 45	35	7.0	161687	2	AC117334	AC117334 Rattus no

ALIGNMENTS

RESULT 1	AC005277	AC005277	118788 bp	DNA	linear	PRI 23-JUL-1998
LOCUS	AC005277	Homo sapiens	chromosome 17,	clone hRPK.597_M_12,	complete sequence.	
DEFINITION	AC005277	Homo sapiens	chromosome 17,	clone hRPK.597_M_12,	complete sequence.	
ACCESSION	AC005277	HTG.				
VERSION	AC005277.1	GI:3337311				
KEYWORDS	HTG.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 118788)					
AUTHORS	Birren,B., Fasmann,K., Linton,L., Nusbaum,C. and Lander,E.					
TITLE	Homo sapiens chromosome 17, clone hRPK.597_M_12					
JOURNAL	Unpublished					

REFERENCE
AUTHORS

2 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 23, 1998 this sequence version replaced gi:3335015.
 All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

FEATURES
source

Location/Qualifiers
 1..118788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="17"
 /chromosome="17"
 /clone.lib="RPC1-11 human BAC library"
 795..884

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region
 /rpt_family="L2"
 complement(4609..4895)
 /rpt_family="MER89"
 repeat_region
 complement(4896..5186)
 /rpt_family="AluSp"
 5219..5479
 /rpt_family="L2"
 complement(5481..5740)
 /rpt_family="MER34"
 repeat_region
 complement(5963..6275)
 /rpt_family="AluSp"
 repeat_region
 complement(6411..6494)
 /rpt_family="MER5A"
 6830..7013
 /rpt_family="MLT1C"
 7014..7318
 /rpt_family="AluYb8"
 repeat_region
 7319..7581
 /rpt_family="MLT1C"
 repeat_region
 7721..8032
 /rpt_family="AluJb"
 repeat_region
 8039..8333
 /rpt_family="AluSq"
 repeat_region
 9431..9460
 /rpt_family="AT-rich"
 9662..9829
 /rpt_family="MIR"
 repeat_region
 complement(11302..11728)
 /rpt_family="MER39"
 repeat_region
 complement(11824..11884)
 /rpt_family="L16C"
 repeat_region
 complement(11928..12108)
 /rpt_family="L16C"
 12453..12725
 /rpt_family="AluJb"
 13377..13785
 /rpt_family="MSTB"
 14118..14153
 /rpt_family="(CA)n"
 repeat_region
 complement(15320..15620)
 /rpt_family="AluJo"
 15683..15759
 /rpt_family="(CA)n"
 repeat_region
 complement(15838..15893)
 /rpt_family="MER91A"
 repeat_region
 complement(15915..16080)
 /rpt_family="MER91A"
 repeat_region
 complement(16474..16544)
 /rpt_family="L2"
 17907..17920
 /rpt_family="L1ME"
 18271..18311
 /rpt_family="L2"
 repeat_region
 complement(18368..18529)
 /rpt_family="AluJb"
 repeat_region
 complement(18530..18560)
 /rpt_family="(CA)n"
 repeat_region
 complement(18561..18681)
 /rpt_family="AluJb"
 18760..19065
 /rpt_family="AluSg"
 complement(19638..19896)
 /rpt_family="MIR"
 repeat_region
 complement(19971..20260)
 /rpt_family="L1ME3A"
 20263..20552
 /rpt_family="MER58B"
 complement(20553..20816)
 /rpt_family="L1ME3A"
 repeat_region
 complement(20821..21069)
 /rpt_family="L1M4"
 repeat_region
 complement(21071..21100)
 /rpt_family="(TAAA)n"

LIHETATLDGEEBAVEKTHSTTSOALSIVGMNNAEFLMLNHFORSORVAKVPLPSPNFS
EKGVADPHMKVCFGDPFTMPKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE
LAGLEDGEPOQKRAHTEEPQAKVRAQ"
BASE COUNT 702 a 793 c 826 g 587 t
ORIGIN

Query Match 60.5%; Score 303; DB 9; Length 2908;
Best Local Similarity 100.0%; Pred. No. 7.2e-93;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 2606 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 2665
Qy 61 AGGACTGGTGTGGCAGCAGCCGGGGCAGGAGGTGCCACACGGAAGCAAGCAGATGA 120
Db 2666 AGGACTGGTGTGGCAGCAGCCGGGGCAGGAGGTGCCACACGGAAGCAAGCAGATGA 2725
Qy 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAGTCTTTGGAACACAGCGGCACCTTTC 180
Db 2726 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAGTCTTTGGAACACAGCGGCACCTTTC 2785
Qy 181 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATGATCAGTG 240
Db 2786 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATGATCAGTG 2845
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGATTTTG 300
Db 2846 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGATTTTG 2905
Qy 301 CAA 303
Db 2906 CAA 2908

RESULT 3
AK074244
LOCUS AK074244 2163 bp mRNA linear PRI 15-FEB-2002
DEFINITION Homo sapiens cDNA FLJ23664 f1s, clone HEP03495, highly similar to
Putative prostate cancer susceptibility protein.
ACCESSION AK074244
VERSION AK074244.1 GI:18676795
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HePG2 cDNA to mRNA, clone_lib:HEP
clone:HEP03495.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2163)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end and one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
Location/Qualifiers
1..2163
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="HEP03495"
/cell_line="HePG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/notes="cloning vector: pME18SFL3"
BASE COUNT 530 a 582 c 613 g 438 t
ORIGIN

Query Match 60.2%; Score 301.4; DB 9; Length 2163;
Best Local Similarity 99.7%; Pred. No. 2.4e-92;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 60
Db 1845 GGTATGGAGCTGTGCGGAGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT 1904
Qy 61 AGGACTGGTGTGGCAGCAGCCGGGGCAGGAGGTGCCACACGGAAGCAAGCAGATGA 120
Db 1905 AGGACTGGTGTGGCAGCAGCCGGGGCAGGAGGTGCCACACGGAAGCAAGCAGATGA 1964
Qy 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAGTCTTTGGAACACAGCGGCACCTTTC 180
Db 1965 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAGTCTTTGGAACACAGCGGCACCTTTC 2024
Qy 181 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATGATCAGTG 240
Db 2025 CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAAAGCAGATGATCAGTG 2084
Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGATTTTG 300
Db 2085 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAGATTGATTTTG 2144
Qy 301 CAA 303
Db 2145 CAA 2147

RESULT 4
BC001939
LOCUS BC001939 2997 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, putative prostate cancer susceptibility protein,
clone MGC:4102 IMAGE:2820640, mRNA, complete cds.
ACCESSION BC001939
VERSION BC001939.1 GI:12804972
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2997)
Strausberg,R.
Direct Submission
Submitted (29-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeeidi, Jacqueline
Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 9 Row: 1 Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7022621.

FEATURES

source
1..2997
/organism="Homo sapiens"
/db_xref="LocusID:60528"
/db_xref="taxon:9606"
/clone="MGC:4102 IMAGE:2820640"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
61..2541
/codon_start=1
/product="putative prostate cancer susceptibility protein"
/protein_id="AAH04158.1"
/db_xref="GI:12804973"
/translation="MMALCSLLRSAAAGRTMSQGTISQAPARRPRPKDPLRLHRLTR
KRGSGCGGNTVYLVVAGSRDGAALVVFSEFNRYLFCNCGEVORLMQEHKLKV
ARLDNIFLTRHNSVGLSGMLTKETGLPKVLSGPPQLEKYLEAIKIFSGPLKG
TELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQSPERPLSRSPERSSELN
ENEPHLPBGVSQRRGVRDSSLVAFICKLHLKRGNFVLKAKEMGLPVGTAAIPIIA
AVKDGKSTHEGREILAEELCTPPDPGAFFVVECPDESFIQICENATFORYOQKAD
APVALVHMAPASVLVDSRYQWMERFGPDTQHLVNLNENCASVHNLRSKHTQTLNLI
HPIIDPLTSFRCKEGTSLVPMVQEGCLLKYQLRPREWORDAITCNPEEFIVEA
LQLPNQSVQYRSDQGPAPAKRSQYPEIIFLGTSALPKIRNVATLVNISP
DTSLLDCGEGTQGLRGHYQDVRLGTLAAVFNHSHADHTGLPSILQRRERL
ASLGLPLHLVAVNQLKQVHQQNCOEVLHHSIMIPAKLQEGAEISPAVERL
ISSLRDCLLEEFOTCLVRHCKHAFGALVHTSGWKVYSGDTMPCCEALVMGKDATL
LIHEATLEDGLEEAEVETHSTTSQASISVGMNNAEFIMLNHFORSYAKVPLFSNFS
EKGVAFDHMKVCFDEPTMPLKIPKLALFAGDIEEMEERREKRELQVRAALLSRE
LAGLEDGEPPQKRAHTEEPQAKKVRQA"

CDS

BASE COUNT 740 a 806 c 853 g 598 t
ORIGIN
Query Match 60.2%; Score 301.4; DB 9; Length 2997;
Best Local Similarity 99.7%; Pred. No. 2.6e-92;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTATGAGCTGTCCGAGCTTGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db 2666 GGTATGAGCTGTCCGAGCTTGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2725
QY 61 AGGACTGGTGGCTGGACACGCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 120
Db 2726 AGGACTGGTGGCTGGACACGCGCGGCGGAGGCTGCCACACGGAAGCAGATGA 2785
QY 121 ACTAATTCATTTCAAGGAGTCTTTTAAAGAGCTCTTTGGAACAGACGCGGACCTTTC 180
Db 2786 ACTAATTCATTTCAAGGAGTCTTTTAAAGAGCTCTTTGGAACAGACGCGGACCTTTC 2845
QY 181 CTCTAATCCAGCAAGTGTATTCCTCGACACACGAGACAGCAGAGTAACAGGATCAGTG 240
Db 2846 CTCTAATCCAGCAAGTGTATTCCTCGACACACGAGACAGCAGAGTAACAGGATCAGTG 2905
QY 241 GGTCTAAGTGTCCGAGACTTAACGAAAAAGTAGTATTTACGTCGCAATAAGATTGAGTTTG 300
Db 2906 GGTCTAAGTGTCCGAGACTTAACGAAAAAGTAGTATTTACGTCGCAATAAGATTGAGTTTG 2965
QY 301 CAA 303
Db 2966 CAA 2968

RESULT 5
BC004158
LOCUS BC004158 3006 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, putative prostate cancer susceptibility protein, clone MGC:2441 IMAGE:2820640, mRNA, complete cds.

ACCESSION

BC004158 GI:13278770

VERSION

MGC.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/BTP

cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadansystemsbio.org

Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia

Greene, Mark Kettman and Anuradha Madan

FEATURES

source

Location/Qualifiers

1..3006

/organism="Homo sapiens"

/db_xref="LocusID:60528"

/db_xref="taxon:9606"

/clone="MGC:2441 IMAGE:2820640"

/tissue_type="Lung, small cell carcinoma"

/lab_host="NIH MGC 7"

/note="Vector: pOTB7"

70..2550

/codon_start=1

/product="putative prostate cancer susceptibility protein"

/protein_id="AAH04158.1"

/db_xref="GI:13278771"

/translation="MMALCSLLRSAAAGRTMSQGTISQAPARRPRPKDPLRLHRLTR KRGSGCGGNTVYLVVAGSRDGAALVVFSEFNRYLFCNCGEVORLMQEHKLKV ARLDNIFLTRHNSVGLSGMLTKETGLPKVLSGPPQLEKYLEAIKIFSGPLKG TELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQSPERPLSRSPERSSELN ENEPHLPBGVSQRRGVRDSSLVAFICKLHLKRGNFVLKAKEMGLPVGTAAIPIIA AVKDGKSTHEGREILAEELCTPPDPGAFFVVECPDESFIQICENATFORYOQKAD APVALVHMAPASVLVDSRYQWMERFGPDTQHLVNLNENCASVHNLRSKHTQTLNLI HPIIDPLTSFRCKEGTSLVPMVQEGCLLKYQLRPREWORDAITCNPEEFIVEA LQLPNQSVQYRSDQGPAPAKRSQYPEIIFLGTSALPKIRNVATLVNISP DTSLLDCGEGTQGLRGHYQDVRLGTLAAVFNHSHADHTGLPSILQRRERL ASLGLPLHLVAVNQLKQVHQQNCOEVLHHSIMIPAKLQEGAEISPAVERL ISSLRDCLLEEFOTCLVRHCKHAFGALVHTSGWKVYSGDTMPCCEALVMGKDATL LIHEATLEDGLEEAEVETHSTTSQASISVGMNNAEFIMLNHFORSYAKVPLFSNFS EKGVAFDHMKVCFDEPTMPLKIPKLALFAGDIEEMEERREKRELQVRAALLSRE LAGLEDGEPPQKRAHTEEPQAKKVRQA"

CDS

70..2550

/product="putative prostate cancer susceptibility protein"

/protein_id="AAH04158.1"

/db_xref="GI:13278771"

/translation="MMALCSLLRSAAAGRTMSQGTISQAPARRPRPKDPLRLHRLTR KRGSGCGGNTVYLVVAGSRDGAALVVFSEFNRYLFCNCGEVORLMQEHKLKV ARLDNIFLTRHNSVGLSGMLTKETGLPKVLSGPPQLEKYLEAIKIFSGPLKG TELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQSPERPLSRSPERSSELN ENEPHLPBGVSQRRGVRDSSLVAFICKLHLKRGNFVLKAKEMGLPVGTAAIPIIA AVKDGKSTHEGREILAEELCTPPDPGAFFVVECPDESFIQICENATFORYOQKAD APVALVHMAPASVLVDSRYQWMERFGPDTQHLVNLNENCASVHNLRSKHTQTLNLI HPIIDPLTSFRCKEGTSLVPMVQEGCLLKYQLRPREWORDAITCNPEEFIVEA LQLPNQSVQYRSDQGPAPAKRSQYPEIIFLGTSALPKIRNVATLVNISP DTSLLDCGEGTQGLRGHYQDVRLGTLAAVFNHSHADHTGLPSILQRRERL ASLGLPLHLVAVNQLKQVHQQNCOEVLHHSIMIPAKLQEGAEISPAVERL ISSLRDCLLEEFOTCLVRHCKHAFGALVHTSGWKVYSGDTMPCCEALVMGKDATL LIHEATLEDGLEEAEVETHSTTSQASISVGMNNAEFIMLNHFORSYAKVPLFSNFS EKGVAFDHMKVCFDEPTMPLKIPKLALFAGDIEEMEERREKRELQVRAALLSRE LAGLEDGEPPQKRAHTEEPQAKKVRQA"

BASE COUNT 742 a 808 c 858 g 598 t

ORIGIN

Query Match

Best Local Similarity

Matches 302; Conservative

0; Mismatches

1; Indels

0; Gaps

0; Score

301.4; DB 9; Length

3006;

Pred. No. 2.6e-92;

QY 1 GGTATGAGCTGTCCGAGCTTGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60

Db 2675 GGTATGAGCTGTCCGAGCTTGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2734

QY 61 AGGACTGTGTCCGAGCTTGGCTCCACATAGCACTAGTCTATAGATGCTCTT 120

```

Db 2735 AGGACTGGTGGCTGGCACACCGCGGGCCAGAGGCTGCCACACGAGCAGAGATGA 2794
QY 121 ACTAATTTCAATTTCAAGGAGTGTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 2795 ACTAATTTCAATTTCAAGGAGTGTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 2854
QY 181 CTCCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGAGAGTAAACAGGATCAGTG 240
Db 2855 CTCCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGAGAGTAAACAGGATCAGTG 2914
QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTCAATAAAGATTTGAGTTTG 300
Db 2915 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTCAATAAAGATTTGAGTTTG 2974
QY 301 CAA 303
Db 2975 CAA 2977

RESULT 6
AK094687
LOCUS
DEFINITION
Hom sapiens prostate cancer susceptibility protein HPC2/ELAC2
mRNA.
ACCESSION
AK094687
VERSION
AK094687.1 GI:21753794
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Hom sapiens
clone:BRAMY2024530.
ORGANISM
Hom sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuko,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
JOURNAL
2 (bases 1 to 2907)
REFERENCE
Isogai,T. and Yamamoto,J.
Direct Submission
TITLE
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
1. 2907
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAMY2024530"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="cloning vector: pME18SFL3"
BASE COUNT
696 a 788 c 839 g 584 t
ORIGIN
Query Match 59.8%; Score 299.4; DB 9; Length 2907;
Best Local Similarity 99.7%; Pred. No. 1.3e-91;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GGTATGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATCCCTCTT 60
Db 2607 GGTATGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATCCCTCTT 2666
QY 61 AGGACTGGTGGCTGGCACAGCGCGGGCCAGAGGCTGCCACACGAGCAAGCAGATGA 120
Db 2667 AGGACTGGTGGCTGGCACAGCGCGGGCCAGAGGCTGCCACACGAGCAAGCAGATGA 2726
QY 121 ACTAATTTCAATTTCAAGGAGTGTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 180
Db 2727 ACTAATTTCAATTTCAAGGAGTGTAAAGAGTCTTGGAAACAGACGCGGCACCTTTC 2786
QY 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGAGTAAACAGGATCAGTG 240
Db 2787 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGAGTAAACAGGATCAGTG 2846
QY 241 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTGCAATAAAGATTTGAGTTTG 300
Db 2847 GGTCTAAGTGTCCGAGACTTAAAGAAATAGTATTTACGCTGCAATAAAGATTTGAGTTTG 2906
QY 301 C 301
Db 2907 C 2907

RESULT 7
AF308698
LOCUS
DEFINITION
Pan troglodytes ELAC2 mRNA, complete cds.
ACCESSION
AF308698
VERSION
AF308698.1 GI:10946496
KEYWORDS
Pan troglodytes.
SOURCE
Pan troglodytes
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (bases 1 to 2908)
AUTHORS
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
JOURNAL
21096977
MEDLINE
11175785
REFERENCE
2 (bases 1 to 2908)
AUTHORS
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Pan troglodytes ortholog of human HPC2/ELAC2
Unpublished
JOURNAL
3 (bases 1 to 2908)
REFERENCE
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Dayananth,P.,
Desrochers,M., Dumont,M., Farnham,J.M., Frank,D., Frye,C.,
Ghaffari,S., Gupte,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J.,
Thomas,A., Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Direct Submission
JOURNAL
Submitted (27-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt

```

FEATURES	Lake City, UT 84103, USA		HYSEQ, INC. (US)	
source	Location/Qualifiers		Location/Qualifiers	
	1..2908		1..2992	
	/organism="Pan troglodytes"		/organism="Homo sapiens"	
	/db_xref="taxon:9598"		/db_xref="taxon:9606"	
CDS	1..2481		75..2555	
	/product="ELAC2"		/note="unnamed protein product"	
	/codon_start=1		/codon_start=1	
	/protein_id="AAG24920.1"		/protein_id="CAD34932.1"	
	/db_xref="GI:10946497"		/db_xref="GI:21439090"	
	/translation="MMALCSLLRSAAAGTMSQRTISOAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCGGVORLMOEHKLKV ARLONIPLTRHMSNVGSLGSMILTKETGLPKVLSPQPOLEKYLEAIKIPSGPLK LELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQWSPERPLSRSPERSDESSEN ENELPHRSGVORRVDSSLVAFICKHLKRGFLVLKAKEMGLPVGTAIAPIIA AVKDGKSTHEGREILABELCTPPDGAAFFVVECPDESFIQICENATFORYOGKAD APVALVHMAESVLVDROYQOMRFGPDTQHLVLNENCASVHNLRSKHTQIOLNLI HPDIPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERQWDAITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSATLVNISP DTSLLDCGEGTFQGLCRHYGDQVDRVLGTLAAVFSVSHLHADHTGLNLTLLQERAL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMPAKCLOEGAEIISPAVERL ISLLRTCDLFEFOTCLVRHCKHAFGCAVHTSGMKVYSGDTMPCEALVRMGKDATL LIHEATLDGLEEEAVENTHSTTSOASVGMNMAEFTMLNHFQRYAKVPLFSPNPN EKVGVAFDHMKVCFDPTMPKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKVRAQ"		/translation="MMALCSLLRSAAAGTMSQRTISOAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCGGVORLMOEHKLKV ARLONIPLTRHMSNVGSLGSMILTKETGLPKVLSPQPOLEKYLEAIKIPSGPLK LELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQWSPERPLSRSPERSDESSEN ENELPHRSGVORRVDSSLVAFICKHLKRGFLVLKAKEMGLPVGTAIAPIIA AVKDGKSTHEGREILABELCTPPDGAAFFVVECPDESFIQICENATFORYOGKAD APVALVHMAESVLVDROYQOMRFGPDTQHLVLNENCASVHNLRSKHTQIOLNLI HPDIPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERQWDAITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSATLVNISP DTSLLDCGEGTFQGLCRHYGDQVDRVLGTLAAVFSVSHLHADHTGLNLTLLQERAL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMPAKCLOEGAEIISPAVERL ISLLRTCDLFEFOTCLVRHCKHAFGCAVHTSGMKVYSGDTMPCEALVRMGKDATL LIHEATLDGLEEEAVENTHSTTSOASVGMNMAEFTMLNHFQRYAKVPLFSPNPN EKVGVAFDHMKVCFDPTMPKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKVRAQ"	
BASE COUNT	712 a	788 c	819 g	589 t
ORIGIN				
	Query Match 59.2%; Score 296.6; DB 9; Length 2908; Best Local Similarity 98.7%; Pred. No. 1.2e-90; Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	60	
Db	2606	GGTATGGAGCTGTGCCAAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	2665	
QY	61	AGGACTGTGCTGGCAGACAGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	120	
Db	2666	AGGACTGTGCTGGCAGACAGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	2725	
QY	121	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGCTTTTGAAGACAGACGCGGCACTTTC	180	
Db	2726	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGGCTTTTGAAGACAGACGCGGCACTTTC	2785	
QY	181	CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAGAGATGAACAGGATCAGTG	240	
Db	2786	CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAGAGATGAACAGGATCAGTG	2845	
QY	241	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	300	
Db	2846	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	2905	
QY	301	CAA 303		
Db	2906	CAA 2908		
RESULT 8				
AX405825				
LOCUS	AX405825	2992 bp	DNA	linear
DEFINITION	Sequence 240 from Patent WO0222660.			
ACCESSION	AX405825			
VERSION	AX405825.1	GI:21439089		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1			
	Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,			
	Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.			
TITLE	Novel nucleic acids and polypeptides			
JOURNAL	Patent: WO 0222660-A 240 21-MAR-2002;			
FEATURES	Lake City, UT 84103, USA		HYSEQ, INC. (US)	
source	Location/Qualifiers		Location/Qualifiers	
	1..2908		1..2992	
	/organism="Pan troglodytes"		/organism="Homo sapiens"	
	/db_xref="taxon:9598"		/db_xref="taxon:9606"	
CDS	1..2481		75..2555	
	/product="ELAC2"		/note="unnamed protein product"	
	/codon_start=1		/codon_start=1	
	/protein_id="AAG24920.1"		/protein_id="CAD34932.1"	
	/db_xref="GI:10946497"		/db_xref="GI:21439090"	
	/translation="MMALCSLLRSAAAGTMSQRTISOAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCGGVORLMOEHKLKV ARLONIPLTRHMSNVGSLGSMILTKETGLPKVLSPQPOLEKYLEAIKIPSGPLK LELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQWSPERPLSRSPERSDESSEN ENELPHRSGVORRVDSSLVAFICKHLKRGFLVLKAKEMGLPVGTAIAPIIA AVKDGKSTHEGREILABELCTPPDGAAFFVVECPDESFIQICENATFORYOGKAD APVALVHMAESVLVDROYQOMRFGPDTQHLVLNENCASVHNLRSKHTQIOLNLI HPDIPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERQWDAITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSATLVNISP DTSLLDCGEGTFQGLCRHYGDQVDRVLGTLAAVFSVSHLHADHTGLNLTLLQERAL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMPAKCLOEGAEIISPAVERL ISLLRTCDLFEFOTCLVRHCKHAFGCAVHTSGMKVYSGDTMPCEALVRMGKDATL LIHEATLDGLEEEAVENTHSTTSOASVGMNMAEFTMLNHFQRYAKVPLFSPNPN EKVGVAFDHMKVCFDPTMPKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKVRAQ"		/translation="MMALCSLLRSAAAGTMSQRTISOAPARRERPRKDPRLHLRTRE KRGPSGSGGNTVYLOVVAAGSRDGAALVVFSEFNRYLFNCGGVORLMOEHKLKV ARLONIPLTRHMSNVGSLGSMILTKETGLPKVLSPQPOLEKYLEAIKIPSGPLK LELAVRPHSAPEYEDMTVYQIPIHSEQRKQHPQWSPERPLSRSPERSDESSEN ENELPHRSGVORRVDSSLVAFICKHLKRGFLVLKAKEMGLPVGTAIAPIIA AVKDGKSTHEGREILABELCTPPDGAAFFVVECPDESFIQICENATFORYOGKAD APVALVHMAESVLVDROYQOMRFGPDTQHLVLNENCASVHNLRSKHTQIOLNLI HPDIPLLTSPCKKEGPTLSVPMVQCECLKYQLRPRERQWDAITCNPEEFIEA LQLENFQOSVQERYRSADGPAPAEKRSQYPEIIFLTGSAIPMKIRNVSATLVNISP DTSLLDCGEGTFQGLCRHYGDQVDRVLGTLAAVFSVSHLHADHTGLNLTLLQERAL ASLGPFLHLLVAPNQLKAWLQYHNCQCEVLHHISMPAKCLOEGAEIISPAVERL ISLLRTCDLFEFOTCLVRHCKHAFGCAVHTSGMKVYSGDTMPCEALVRMGKDATL LIHEATLDGLEEEAVENTHSTTSOASVGMNMAEFTMLNHFQRYAKVPLFSPNPN EKVGVAFDHMKVCFDPTMPKLIPLKALFAGDIEEMERERREKRELQVRAALLSRE LAGGLEDEPQOKRAHTEEPQAKVRAQ"	
BASE COUNT	712 a	788 c	819 g	589 t
ORIGIN				
	Query Match 59.2%; Score 296.6; DB 6; Length 2992; Best Local Similarity 98.7%; Pred. No. 1.2e-90; Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	60	
Db	2606	GGTATGGAGCTGTGCCAAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCTCTT	2739	
QY	61	AGGACTGTGCTGGCAGACAGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	120	
Db	2740	AGGACTGTGCTGGCAGACAGCGCGGCGAGAGGCTGCCACAGCAAGCAGAGATGA	2799	
QY	121	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGCTTTTGAAGACAGACGCGGCACTTTC	180	
Db	2800	ACTAATTTTCATTTCAAGGAGCTTTTAAAGAGCTTTTGAAGACAGACGCGGCACTTTC	2859	
QY	181	CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAGAGATGAACAGGATCAGTG	240	
Db	2860	CTCTAATCCAGCAAAAGTGATTCCTCGCACACAGAGACAGAGATGAACAGGATCAGTG	2919	
QY	241	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	300	
Db	2920	GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAGATTGAGTTTG	2979	
QY	301	CAA 303		
Db	2980	CAA 2982		
RESULT 9				
AK094333				
LOCUS	AK094333	2734 bp	mRNA	linear
DEFINITION	Homo sapiens cDNA FLJ37014 fis, clone BRACE2010203, highly similar			
KEYWORDS	to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2			
SOURCE	clone:BRACE2010203.			
ACCESSION	AK094333			
VERSION	AK094333.1	GI:21753374		
KEYWORDS	oligo capping; fis (full insert sequence).			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1			
	Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,			

Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2734)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2010203"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"

BASE COUNT 671 a 708 c 724 g 631 t

ORIGIN
Query Match 58.8%; Score 294.6; DB 9; Length 2734;
Best Local Similarity 98.7%; Pred. No. 5.7e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db 2434 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2493

Qy 61 AGGACTGTGCTGGCAGCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2494 AGGACTGTGCTGGCAGCGCTGGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2553

Qy 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTGGAAACACAGCGCGGCACCTTTC 180
Db 2554 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTGGAAACACAGCGCGGCACCTTTC 2613

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGCAGATGAACAGGATCAGTG 240
Db 2614 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGCAGATGAACAGGATCAGTG 2673

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 300
Db 2674 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 2733

Qy 301 C 301
Db 2734 C 2734

RESULT 10
AK094012
LOCUS
DEFINITION Homo sapiens cDNA FLJ36693 fis, clone UTERU2008901, highly similar to Homo sapiens putative prostate cancer susceptibility protein HPC2/ELAC2 mRNA.
ACCESSION AK094012
VERSION AK094012.1 GI:21752986
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens uterus cDNA to mRNA, clone_lib:UTERU2 clone:UTERU2008901.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiya,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2793)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UTERU2008901"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

BASE COUNT 610 a 754 c 806 g 623 t

ORIGIN
Query Match 58.8%; Score 294.6; DB 9; Length 2793;
Best Local Similarity 98.7%; Pred. No. 5.7e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 60
Db 2493 GGTATGGAGCTGTGCGAGGCTTGGGCTCCACATAGCACTAGTCTATAGATGCTCTT 2552

Qy 61 AGGACTGTGCTGGCAGCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db 2553 AGGACTGTGCTGGCAGCGCTGGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 2612

Qy 121 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTGGAAACACAGCGCGGCACCTTTC 180
Db 2613 ACTAATTTTCATTTCAAGGCAGTTTTTAAAGAAAGTCTTGGAAACACAGCGCGGCACCTTTC 2672

Qy 181 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGCAGATGAACAGGATCAGTG 240
Db 2673 CTCTAATCCAGCAAGTGTATCCCTGCGACACAGAGACAGCAGATGAACAGGATCAGTG 2732

Qy 241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 300
Db 2733 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGTGTCAATAAGATTGAGTTTG 2792

Qy 301 C 301
Db 2793 C 2793

RESULT 11
AK001392
LOCUS
DEFINITION Homo sapiens cDNA FLJ10530 fis, clone NT2RP2000985.
ACCESSION AK001392

```
VERSION AK001392.1 GI:7022621
KEYWORDS oligo capping; fits (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,
        clone_lib:NT2RP2 clone:NT2RP2000985.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
        Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
        Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
        Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
        Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
        Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
        Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2976)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
        Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
        (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
        International Trade and Industry of Japan; cDNA full insert
        sequencing: Research Association for Biotechnology; cDNA library
        construction, 5'- & 3'-end one pass sequencing and clone selection:
        Helix Research Institute (supported by Japan Key Technology Center
        etc.) and Department of Virology, Institute of Medical Science,
        University of Tokyo.
FEATURES             source
SOURCE              1..2976
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="NT2RP2000985"
                    /cell_line="NT2"
                    /cell_type="teratocarcinoma"
                    /clone_lib="NT2RP2"
                    /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
                    precursor cells after 2-weeks retinoic acid (RA)
                    induction."
CDS
75..2639
     /note="unnamed protein product"
     /codon_start=1
     /protein_id="BAA91666.1"
     /db_xref="GI:7022622"
     /translation="MWALCLLSAAGRTMSQRTISOAPARRRPRKDLRLHLRTRE
KEGSCSGGPNVTVLQVVAAGSRDGAALYVSEFNRYLNCGEVORLQEHKLK
ARLDNIFTRMWSNVGSLGSMILTKETGLPKCVSPPPQLEKYLEAIKIFSGPLK
IELAMRHPASPEYEDETVTYQIPIHSEQRGKHQWQSPERPLSRSPSSSEN
ENEPHLPVGSORRGRVDSLSVAFICKLHKGNFVLKAKEMGLPVGTAIPIIA
AVDKGSIITHEGREILABELCTPPDPAFVVVECPDESFTQICENATFORQOGKAD
APVALVHMASVLVDSRYOQMERFGPDTOHLVNLNENASVHNLRSKHQIOTLNLI
HPDIFPLTSPCKEGPTLSVPMVQEGCLLKQLRPREWQDAILTCNPEFVEAL
QUPNFQSVETRRRAQGPAPAEKRSQPEIFIGTSAIPMKIRNVSAITLVNISPD
TSLDDQCEGTGQCRHYGDQVDRVLGTLAAVFVSHLHADHTGLPSILIQERALA
SLGKLEPILVLPVAPLQKQYNNQOEVLHISMTIPAKCLOEAGISIPAVERLI
SLLRTCDLEEFQTCVLRKCHAFGCVLHVSGWVYSGDTPCEALVRMGKDATLL
THEATLEDLEEEAVEKTHSTTSQAI SVCMRNNAEFIMLHFESQRYAKVPLFSPNSE
KGVAFDHKVCGLGFPPTMPKLIPLKALFAGDIEEMERREKRELQVRAALLSREL
AGGLSDGSLRSRSPGTPRRSPSPSDGLGDPFLRLRLVCFCPTHAPVSALLAGRS"
BASE COUNT      712 a      807 c      856 g      601 t
ORIGIN
Query Match      58.8%; Score 294.6; DB 9; Length 2976;
Best Local Similarity 98.7%; Pred. No. 5.8e-90;
Matches 297; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 60
      |||||||
Db 2676 GGTATGGAGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGTCTATAGATGCCTCTT 2735
      |||||||
QY 61 AGGACTGGTCTGGCACACGCCCGCGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGA 120
      |||||||
```


/clone_lib="Human"	
STS	106..427
primer_bind	106..128
primer_bind	complement (405..427)
BASE COUNT	91 a 115 c 108 g 115 t
ORIGIN	
Query Match	57.2%; Score 286.4; DB 11; Length 429;
Best Local Similarity	99.3%; Pred. No. 2.4e-87;
Matches 298; Conservative	0; Mismatches 1; Indels 1; Gaps 1;
QY	1 GGTATGGAGCTGTGCGAGGCTTGGCTCCACATAAGCAGCTAGTCTATAGATGCTCTT 60
Db	299 GGTATGGAGCTGTGCGAGGCTTGGCTCCACATAAGCAGCTAGTCTATAGATGCTCTT 240
QY	61 AGGACTGTGCTGGCCACACGCCGGCGAGGCTGCCACACGGAAGCAAGCAGATGA 120
Db	239 AGGACTGTGCTGCT-GCACACGCCGGCGAGGCTGCCACACGGAAGCAAGCAGATGA 181
QY	121 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTTGTGAAACACACGCCGACCTTTC 180
Db	180 ACTAATTTTCATTTCAAGGAGCTTTTAAAGAAGCTATGGAACACACGCCGCGACCTTTC 121
QY	181 CTCTAATCAGCAAGTAGTATCCCTGCACACACAGAGACAAGCAGAGTAACAGGATCAGTG 240
Db	120 CTCTAATCAGCAAGTAGTATCCCTGCACACACAGAGACAAGCAGAGTAACAGGATCAGTG 61
QY	241 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTG 300
Db	60 GGTCTAAGTGTCCGAGACTTAACGAAATAGTATTTCAGCTGCAATAAAGATTGAGTTG 1
RESULT 13	
AF308694	
LOCUS	Gorilla gorilla ELAC2 (ELAC2) mRNA, complete cds.
DEFINITION	
ACCESSION	AF308694
VERSION	AF308694.1 GI:10946488
KEYWORDS	
SOURCE	Gorilla gorilla.
ORGANISM	Gorilla gorilla.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
AUTHORS	1 (bases 1 to 2893) Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliiev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J., and Cannon-Albright,L.A.
TITLE	A candidate prostate cancer susceptibility gene at chromosome 17p
JOURNAL	Nat. Genet. 27 (2), 172-180 (2001)
MEDLINE	21096977
PUBMED	11175785
REFERENCE	2 (bases 1 to 2893) Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliiev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J., and Cannon-Albright,L.A.
AUTHORS	2 (bases 1 to 2893) Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliiev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J., and Cannon-Albright,L.A.
TITLE	Gorilla gorilla ortholog of human HPC2/ELAC2
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2893) Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliiev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J., and Cannon-Albright,L.A.
AUTHORS	3 (bases 1 to 2893) Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M., Dumont,M., Farnham,J.M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S., Gupte,J.S., Hu,R., Iliiev,D., Janacki,T., Kort,E.N., Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R., Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M., Woodland,A.M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J., and Cannon-Albright,L.A.

STS size: 137 PCR Profile:	Presoak: 94 degrees C for 4.00 minutes Denaturation: 94 degrees C for 50.0 seconds Annealing: 58 degrees C for 1.50 minutes Polymerization: 72 degrees C for 1.00 minutes PCR Cycles: 30 Thermal Cycler: custom built by IAS, Costar, Cambridge MA
Protocol:	Template: 10 ng Primer: each 5 pM dNTPs: 4 mM Taq Polymerase: 0.5 U Total Vol: 20 uL
Buffer:	Mg2+: 1.5 mM KCl: 50 mM Tris-HCl: 10 mM Gelatin: .001 %
FEATURES	Location/Qualifiers
source	1..137 /organism="Homo sapiens" /db_xref="taxon:9606" /map="36.40 cR from top of Chr17 linkage group" /clone_lib="Human Thudson EST" /notes="STSS derived from sequences in dbEST and the Unigene collection."
STS	1..137
primer_bind	1..25
primer_bind	complement(118..137)
BASE COUNT	28 a 30 c 31 g 46 t 2 others
ORIGIN	
Query Match	26.6%; Score 133.4; DB 11; Length 137; Best Local Similarity 97.8%; Pred. No. 1.5e-34; Matches 134; Conservative 0; Mismatches 3; Indels 0; Gaps
Qy 151	AAGTCTTTGGAAACAGACGGCGGACCTTTCTCTTAATCCAGCAAAGTGATTCCTCGACA 210
Db	
137	AAGTCATGGAAACAGACGGCGGACCTTTCTCTTAATCCAGCAAANTGATTCCTCGANA 78
Qy 211	CCAGAGACAAGCAGAGTAAACAGGATCAGTGGGTCTTAAGTCTCCGAGACTTAAACGAAAATA 270
Db	
77	CCAGAGACAAGCAGAGTAAACAGGATCAGTGGGTCTTAAGTCTCCGAGACTTAAACGAAAATA 18
Qy 271	GTATTTTCAGTGCATA 287
Db	
17	GTATTTTCAGTGCATA 1
Search completed:	May 17, 2003, 21:41:54
Job time :	1383.06 secs

AX069570	GI:12579355
VERSION	
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J., Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D., Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B. Human genes and gene expression products Patent: WO 0102568-A 42 11-JAN-2001; CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
TITLE	
JOURNAL	
FEATURES	Location/Qualifiers
source	1..386 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	82 a 90 c 111 g 103 t
ORIGIN	
Query Match	31.5%; Score 158; DB 6; Length 386; Best Local Similarity 100.0%; Pred. No. 6.3e-43; Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 344	AGGTCGTGCTGCACCACTTGGAGAAGCTCTCTGTGCTGTAGTGTGGCAGCTGCCGTG 403
Db	
7	AGGTCGTGCTGCACCACTTGGAGAAGCTCTCTGTGCTGTAGTGTGGCAGCTGCCGTG 66
Qy 404	GTACCCGGTGGCTTGGAGAAGTCACTCCGCTGTGTAGTGACACCTCTGGAACCTGTC 463
Db	
67	GTACCCGGTGGCTTGGAGAAGTCACTCCGCTGTGTAGTGACACCTCTGGAACCTGTC 126
Qy 464	CTCAGAGGCACCTTATTTCGCCAGTCTTTTGACA 501
Db	
127	CTCAGAGGCACCTTATTTCGCCAGTCTTTTGACA 164

RESULT 15	
G42927/c	
LOCUS	G42927 137 bp mRNA linear STS 27-JAN-1999
DEFINITION	WIAF-176-STS Human Thudson EST Homo sapiens STS cDNA, sequence tagged site.
ACCESSION	G42927
VERSION	G42927.1 GI:4191844
KEYWORDS	STS.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Wang,D.G., Fan,J.B., Siao,C.J., Berno,A., Young,P., Sapolsky,R., Ghandour,G., Perkins,N., Winchester,E., Spencer,J., Krugiyak,L., Stein,L., Hsie,L., Topaloglou,T., Hubbell,E., Robinson,E., Mittmann,M., Morris,M.S., Shen,N., Kilburn,D., Rioux,J., Nusbaum,C., Rozen,S., Hudson,T.J., Lipshutz,R., Chee,M. and Lander,E.S.
TITLE	Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome
JOURNAL	Science 280 (5366), 1077-1082 (1998)
MEDLINE	98248615
PUBMED	9582121
COMMENT	Synonyms: EST226740b, EST226740 Contact: Thomas Hudson Whitehead Institute/MIT Center for Genome Research Whitehead Institute for Biomedical Research 9 Cambridge Center, Cambridge MA 02142 USA Tel: 617 252 1900 Fax: 617 252 1902 Email: thudson@genome.wi.mit.edu Primer A: TATTGACGTGAATACTATTTCG Primer B: AAGTCATGAAACAGACGGC

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 680.481 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 tatcagtgactgaattctta.....aagctctgagggaactgacgt 500

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_nus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	122.4	24.5	992	12	BF244530
2	119.6	23.9	546	13	BG943515
3	118.2	23.6	669	17	AG177687
4	118.2	23.6	811	17	AQ742365
C 5	117.8	23.6	399	17	AQ587593
C 6	117.8	23.6	450	17	AQ587553

C 7	117.8	23.6	727	9	AU121896
C 8	117.8	23.6	770	17	B02308
C 9	117.6	23.5	415	9	AI609972
C 10	117.6	23.5	669	17	AQ196057
C 11	117.4	23.5	389	17	AQ071874
C 12	116.8	23.4	654	17	AG084157
C 13	116.6	23.3	312	17	AQ230437
C 14	116.4	23.3	389	17	B95260
C 15	116.4	23.3	418	17	AQ0628850
C 16	116.4	23.3	450	17	AQ086239
C 17	116	23.2	527	17	AQ183714
C 18	116	23.2	664	17	AQ343449
C 19	116	23.2	801	17	AQ740452
C 20	115.8	23.2	523	17	AQ677233
C 21	115.6	23.1	398	9	AL696113
C 22	115.6	23.1	500	17	B44892
C 23	115.4	23.1	324	9	AL707132
C 24	115.4	23.1	370	17	AZ756633
C 25	115.4	23.1	439	17	AQ378798
C 26	115.2	23.0	383	9	AA584125
C 27	115.2	23.0	426	9	AL043719
C 28	115.2	23.0	1080	13	BM558580
C 29	115	23.0	421	14	BQ614109
C 30	115	23.0	458	12	BF912236
C 31	114.8	23.0	390	17	AQ096225
C 32	114.6	22.9	670	17	AG013922
C 33	114.4	22.9	339	9	AI820920
C 34	114.4	22.9	426	12	BF589079
C 35	114.4	22.9	443	9	AI821670
C 36	114.4	22.9	450	17	AQ144748
C 37	114.4	22.9	459	9	AA478209
C 38	114.4	22.9	515	17	AQ019249
C 39	114.4	22.9	586	17	AQ583832
C 40	114.4	22.9	667	17	AG059714
C 41	114.4	22.9	896	12	BF342636
C 42	114.4	22.8	516	10	BE148994
C 43	114.2	22.8	579	14	BM990962
C 44	114.2	22.8	600	10	AV759518
C 45	114.2	22.8	691	17	AG047439

ALIGNMENTS

RESULT 1
BF244530/c
LOCUS BF244530 992 bp mRNA linear EST 14-NOV-2000
DEFINITION 601862835f1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4080393 5',
mRNA sequence.
ACCESSION BF244530
VERSION BF244530.1 GI:11158461
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 992)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM940 row: f column: 10
High quality sequence stop: 328.
Location/Qualifiers
1. .992

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4080393"
/clone_lib="NIH_MGC_57"
/tissue_type="glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccattatgcc); Site_2: SfiI (ggccattatgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-3' (30) BN-3, (where B = A, C, G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
BASE COUNT      354 a 259 c 286 g 93 t
ORIGIN

Query Match      24.5%; Score 122.4; DB 12; Length 992;
Best Local Similarity 84.1%; Pred. No. 1e-13;
Matches 138; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 181 TTTTGTGTTTGTGTTTGTGTTTGTGACAGGGTCTCGAGGTGTACCCAGGCTGGAGT 240
      |||||
Db 309 TTTTGTGTTTGTGTTTGTGTTTGTGACAGAGTCTCGTCTGTGCGCGAGCTGGAGT 250
      |||||

QY 241 GCAGTGGCGGATTCGACTCACCGCAACCTCCGCTCCGCGCTTAAGCGATTTCCTGCG 300
      |||||
Db 249 GCAGTGGCGGATTCGCTCACTCAACCTCTGCTCCGAGTTCAGCGATTCTCTCTCG 190
      |||||

QY 301 CTCAGCTCCCAAGTAGCTGGGACTACAGCTCGGGACACACACG 344
      |||||
Db 189 CTCAGCTCTGAGTAGCTGGGATTACAGGCGCGCGCCACACG 146
      |||||

RESULT 2
BG943515 546 bp mRNA linear EST 11-JUN-2001
LOCUS ax38g05.x1 Proliferating Human Erythroid Cells (LCB:ax library)
DEFINITION Homo sapiens cDNA clone ax38g05 random, mRNA sequence.
ACCESSION BG943515
VERSION BG943515.1 GI:14342887
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 546)
AUTHORS Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
TITLE Gene expression in proliferating human erythroid cells
JOURNAL Genomics 59 (2), 168-177 (1999)
MEDLINE 99339981
COMMENT Contact: Jeffery L. Miller
        Laboratory of Chemical Biology
        National Institute of Diabetes and Digestive and Kidney Diseases
        Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
        20892, USA
        Tel: 301 402 2373
        Fax: 301 435 5148
        Email: jmf@nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 38 row: g column: 05
Seq primer: -21M13 forward primer (ABI).
          Location/Qualifiers
            1..546
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="ax38g05"
              /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
              library)"

```

```

/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev stage="Progenitor; EPO responsive CD71+++"
/lab_host="SOLR"
/notes="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI; Site_2: EcoRI; 65,000 proliferating erythroid cells from the Buffy coat of a blood donation were obtained by flow cytometric separation after a 5-day culture period in the presence of erythropoietin. Total RNA was purified from the sorted cell population using TRIzol reagent. RNA (0.3 ug) was converted into double stranded cDNA using Clontech's Capfinder cDNA Library Construction Kit (Clontech) according to the manufacturer's protocol and cloned into EcoRI digested Lambda Zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH intramural sequencing center http://www.nisc.nih.gov/)."
BASE COUNT      136 a 110 c 98 g 202 t
ORIGIN

Query Match      23.9%; Score 119.6; DB 13; Length 546;
Best Local Similarity 80.5%; Pred. No. 4.3e-13;
Matches 140; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 181 TTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGCTCTCGAGGTGTACCCAGGCTGGAGT 240
      |||||
Db 7 TTTTGTGTTTGTGTTTGTGTTTGTGAGACAGTCTCGCTCTGTCGCGCAGGCTGGAGT 66
      |||||

QY 241 GCAGTGGCGGATTCGACTCACCGCAACCTCCGCTCCGCGCTTAAGCGATTTCCTGCG 300
      |||||
Db 67 GCAGTGGCGGATTCGCGCTCACTGCAACCTCTGCTCCGCGGCTCAAGCAATTCTCTGCG 126
      |||||

QY 301 CTCAGCTCCCAAGTAGCTGGGACTCACAGCTCGGGACACACGTTAAATGAT 354
      |||||
Db 127 CTCGGCTCCCGAGTAGCTGGGATTACACAGGTGTGCCACCATGCCCACTAAT 180
      |||||

RESULT 3
AG177687 669 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-049G22.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG177687
VERSION AG177687.1 GI:16707367
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee Male BAC Library clone:RP43-049G22.TJ.
ORGANISM Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 669)
          Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:chimpbes@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
          end was generated during the R&D process and may have higher chance
          of clone tracking errors.
          PRIMERS

```

```

Sequencing: TJ
LIBRARY
  Vector      : pBACE3.6
  R-site 1    : EcoRI
  R-site 2    : EcoRI
  Location/Qualifiers
    1..669
    /organism="Pan troglodytes"
    /db_xref="taxon:9598"
    /clone="RP43-049G22.TU"
    /sex="male"
    /cell_type="lymphocytes"
    /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT 160 a 146 c 130 g 233 t
ORIGIN

Query Match      23.6%; Score 118.2; DB 17; Length 669;
Best Local Similarity 75.4%; Pred. No. 7.4e-13;
Matches 147; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 150 TTCTTGCTTTCTTCCAAAACACTACAAATTTTGTGTTTGTGTTTGTGTTGAGA 209
      |||||
Db 117 TTCTTGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTGAGA 176
      |||||

QY 210 CAGGCTCGAGGTGTCACCCAGGCTGGAGTGCAGTGGCGGATTCGACTCACCGCAAC 269
      |||||
Db 177 TCGAGTCTCACTCTGTCTCCAGGCTGGAGTGCAGTGGCGGATTCGACTCACCGCAAC 236
      |||||

QY 270 CTCGGCTCCGGCTTAAGCGATTCCTCGCTCAGCTCCCAAGTAGCTGGGACTACAA 329
      |||||
Db 237 CTCGGCTCCGGGTTACCGCCATTCCTCGCTCAGCTCCCGATTCGATTCGAGTGGGACTACAG 296
      |||||

QY 330 GTCGGGACACCAACG 344
      |||||
Db 297 GCACCTGCCACCAACG 311
      |||||

RESULT 4
A0742365
LOCUS HS_5382_B1_B08_T7A_RPCI-11_Human_Male_BAC_Library_Homo_sapiens
DEFINITION genomic_clone_Plate-958_Col=15_Row=D, DNA sequence.
ACCESSION A0742365
VERSION A0742365.1 GI:5519887
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(peter@redjones.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 958 row: D column: 15
Seq primer: T7
Class: BAC ends

Sequencing: TJ
LIBRARY
  Vector      : pBACE3.6
  R-site 1    : EcoRI
  R-site 2    : EcoRI
  Location/Qualifiers
    1..811
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="Plate-958_Col=15_Row=D"
    /clone_lib="RPCI-11_Human_Male_BAC_Library"
    /sex="male"
    /note="vector: pBACE3.6; Site1: EcoRI; Site2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"
BASE COUNT 166 a 184 c 170 g 248 t 43 others
ORIGIN

Query Match      23.6%; Score 118.2; DB 17; Length 811;
Best Local Similarity 70.0%; Pred. No. 6.9e-13;
Matches 173; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 131 CTTTAGAGCTTGCTCTATTCCTGCTTTCTTTCCTTCCAAAAACACTACAAATTTTGT 190
      |||||
Db 214 CTGTAATATTGCTTTATTGTTCTATTGTTTTCAGTCATTCGCTGTAATTTTCTTTT 273
      |||||

QY 191 GTTTTGTGTTTGTGAGACAGGTTCTGAGTGTCAACCCAGGCTGGAGTGCAGTGGCG 250
      |||||
Db 274 TTTTGTGTTTGTGAGACAGGTTCTGCTCTGTCTCCAGGCTGGAGTGCAGTGGCG 333
      |||||

QY 251 GATTTTCGACTACCCCAACCTCCGCTCCGCGCTT-AAGCGATTCCTCGCTCAGCCTC 309
      |||||
Db 334 GATCTCGGCTCACTCGAAGCTCCGCTCCGGAATTCACGCCATTCCTCGCTCAGCCTC 393
      |||||

QY 310 CCAAGTAGCTGGGACTACAAGCTCGGGACACCACTGTAATAATGATCAAGTTCTTAACATGT 369
      |||||
Db 394 CGGTGTAGCTGGGACTACAGGGCGCCGCCACCACTGTAATAATGATCAAGTTCTTAACATGT 453
      |||||

QY 370 ATGCATA 376
      |||||
Db 454 AAGTAGA 460
      |||||

RESULT 5
A0587593
LOCUS CITBI-EI-2643H8.TR_CITBI-EI_Homo_sapiens_genomic_clone_2643H8, DNA
DEFINITION sequence.
ACCESSION A0587593
VERSION A0587593.1 GI:5014273
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-EI-2643H8.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hne@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers

```

```
source
1. .399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2643H8"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 119 a 83 c 108 g 89 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 17; Length 399;
Best Local Similarity 73.8%; Pred. No. 1e-12;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATCTTCTTTTCCAAAACACATACAAATTTT 184
Db 399 CTCCTCCCTTCAAATATCAACATTTCTACAATAGAGGCATACAGAGATGACAACCTTTT 340

Qy 185 TGTCTTTGTTTGTGTTTGTGAGACAGGGTCTCAGAGTGTCAACCCAGGCTGGAGTGCAG 244
Db 339 TTTTCTTTTCTTTTCTTTTGTGAGATGAGTCTTCTCTGTCAACCCAGGCTGGAGTGCAG 280

Qy 245 TGGCGCGATTTTCGACTCAGCGAACCTCGGCTCCGCGCTT-AAGCGATTCTCTGCGCTC 303
Db 279 TGGCGCGATCTCGGCTCACTGCAAGTCCGCTCCAGGTTCCAGCCATTCTCTGCGCTC 220

Qy 304 AGCTCCCAAGTAGCTGGGACTACAAAGTCTGGGACACCAAG 344
Db 219 AGCTCCCAAGTAGCTGGGACTACAGGACCCGCCCATG 179

RESULT 6
AQ587553/c
LOCUS
DEFINITION AQ587553 Homo sapiens genomic clone 2643H14,
GSS 07-JUN-1999
CITBI-E1-2643H14.1 TR CITBI-E1 Homo sapiens genomic clone 2643H14,
DNA sequence.
ACCESSION AQ587553
VERSION AQ587553.1 GI:5014233
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Other GSSs: CITBI-E1-2643H14.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1. .450
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2643H14"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 119 a 83 c 108 g 89 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 9; Length 727;
Best Local Similarity 73.8%; Pred. No. 8.6e-13;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATCTTCTTTTCCAAAACACATACAAATTTT 184
Db 312 CTTTCTGTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 371

source
1. .727
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1001220"
/clone_lib="MAMMA1"
/tissue_type="mammary gland"
/notes="Vector: pME18SFL3"
BASE COUNT 170 a 159 c 146 g 248 t
ORIGIN

Query Match 23.6%; Score 117.8; DB 9; Length 727;
Best Local Similarity 73.8%; Pred. No. 8.6e-13;
Matches 163; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

Qy 125 CTCCTGCTTTAGAGCTTGCTCTCTATCTTCTTTTCCAAAACACATACAAATTTT 184
Db 312 CTTTCTGTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 371
```

Qy	185	TGTTTGTGTTTGTGTTTGTGAGACAGGGTCTCGAGGTGTCAACCCAGGCTGGAGTGCAG	244
Db	372	TGTTGTTTGTGTTTGTGTTTGTGAGACAGAGTCTCTCTGTGTCAACCCAGGCTGGAGTGCAG	431
Qy	245	TGGCGGATTTTCGACTCACCCCAACTCCGCGCTTCGGCGCTT-AAAGCGATTCTCTCGCTC	303
Db	432	TGAGCGATCTCGGCTACTGTCAAGTCCGCGCTCCAGGTTCAAGCGGTTTCTTCTGCTC	491
Qy	304	AGCTCCCAAGTAGCTGGAGCTACAAGCTCGGACACCAACG	344
Db	492	AGCTCCCAAGTAGCTGGAGCTACAGGCACCGCCACCAACG	532
RESULT 8			
B02308/c			
LOCUS			
DEFINITION			
B02308 770 bp DNA linear GSS 13-JUL-1996			
csRL-151D1-u csRL flow sorted Chromosome 11 specific cosmid Homo			
sapiens genomic clone csRL-151D1, DNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS			
Evans, G.A., Burbee, D., Davies, C., Mahner, L., Oliver, T., Gilbert, M.,			
Jones, D., Ward, T., Gillilan, B., Schagemann, J., Probst, S., Harris			
J., Deford, J., McFarland, J., Burzinski, K., Khan, M., Kupfer, K. and			
Garner, H.R.			
Genomic Sequence Sampled Map of Chromosome 11			
TITLE			
JOURNAL			
COMMENT			
Unpublished (1996)			
Contact: Evans GA, Shane Probst			
McDermott Center for Human Growth and Development			
University of Texas Southwestern Medical Center At Dallas			
5323 Harry Hines Blvd, Dallas TX 75235-8591			
Tel: 214-648-1600			
Fax: 214-648-1666			
Email: ge Evans@utsw.swmed.edu, shane@mcdermott.swmed.edu			
PCR Primers			
FORWARD: ATAAAGCAGGTTCTTAAG			
BACKWARD: CTTTCTACAGGAAGGAC			
Seq primer: T7			
Class: cosmid ends			
High quality sequence stop: 770.			
Location/Qualifiers			
1. .770			
/organism="Homo sapiens"			
/db xref="taxon:9606"			
/clone="csRL-151D1"			
/clone_lib="csRL flow sorted Chromosome 11 specific			
cosmid"			
/sex="female"			
/cell_type="chimeric hamster somatic cell hybrid"			
/note="vector: scos-1; Human Chromosome 11 specific cosmid			
library prepared from flow sorted human Chromosome 11			
derived from Chinese Hamster Ovary (CHO) monochromosomal			
somatic cell hybrid, J1"			
BASE COUNT			
ORIGIN			
219 a 169 c 185 t 31 others			
Query Match			
Best Local Similarity			
Matches 134; Conservative 0; Mismatches 27; Indels 0; Gaps 0;			
23.6%; Score 117.8; DB 17; Length 770;			
83.2%; Pred No. 8.4e-13;			
Qy 184 TTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGGTGTCAACCCAGGCTGGAGTGC			
Db 270 TTTCTTTTTTTTTTTTTTTTTTGGACAGAGTCTCGCTCTGTCGCGGAGGCTGGAGTGC			
Qy 244 GTGGCGGATTTCGACTCAACCGCAACTCCGCTCCGGCTTAAGCGATTCTCTGCTC			
Db 210 GTGGCGGATCTCGGCTCACTGCAACTCTGCTCCGAGTTCAACGCAATTCTCTGCTC			

Frequency: 2117
T.T.B.B.A.V.

R.Site 1 : SacI

Location/Quali

/organism="Pan

ovary = 2800
/clone="PTB-08

```
/sex="male"  
/cell type="1" w
```

209 a 14E c 1
/clone_lib="PI

23.48%

Conservative 0;

TTTCTTTC AAAA CACT.

CTGCGATACGTACAATTATTA

CGAGGTGTACCCAGGCTG

TGCTCTGTACCCAGGCTG

CCGCGCTTAAGCGATTCTC

CCGGGTTCAAGCAATTCTC

343 CACAC

IIIIIIII
TACCAC 157

0437

ens genomic clone p

0437.1 GI:3655666

n.

ryota; Metazoa; Cho

cases 1 to 312)

er, A., Shaker, R., F

ence-tagged connect

Natl Acad Sci

0589
act: Mahiyas CG W

Throughput Sequenc

Queen Anne Avenue N
(200) 512-3510

(206) 616-3887

ence Tagged Connect

s: BAC ends

Location/Quali

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 17:27:39 ; Search time 63.3902 Seconds
(without alignments)
10174.595 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500
Perfect score: 500
Sequence: 1 tatcagtgactgaattctt.....aagctctgagggaactgaactg 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	26664	9	US-09-988-626-28
2	500	100.0	26664	9	US-09-988-687-28
3	120.4	24.1	7680	10	US-09-939-581A-3
4	120	24.0	1503841	9	US-09-946-807-1
5	120	24.0	1503841	10	US-09-795-668-1
6	120	24.0	1503841	10	US-09-795-686-1
7	119.8	24.0	4045	9	US-09-764-891-8718
8	119.8	24.0	4045	9	US-09-764-891-8719
9	119.6	23.9	1160	8	US-08-927-933-78
10	119.6	23.9	1160	9	US-09-954-531-146
11	118.8	23.8	73467	9	US-10-237-859-3
12	118.4	23.7	17752	10	US-09-748-127-3
13	118.4	23.7	465237	10	US-09-933-267A-1
14	118.2	23.6	5159	10	US-09-764-877-3707
15	118.2	23.6	6834	9	US-09-764-891-8002
16	118.2	23.6	6834	9	US-10-091-438-263
17	118	23.6	1115	9	US-09-764-872-943
18	118	23.6	1115	9	US-09-764-872-944
19	117.6	23.5	6186	9	US-10-074-095-1075

20	117.6	23.5	6186	10	US-09-764-860-1075
21	117.2	23.4	433	9	US-09-946-807-1404
22	117.2	23.4	433	10	US-09-795-668-1404
23	117.2	23.4	433	10	US-09-795-668-1404
24	116.8	23.4	2295	9	US-09-764-891-7825
25	116.4	23.3	5426	10	US-09-798-029-7
26	116	23.2	6191	9	US-10-074-095-1076
27	116	23.2	6191	9	US-10-074-095-1077
28	116	23.2	6191	10	US-09-764-860-1076
29	116	23.2	6191	10	US-09-764-860-1077
30	116	23.2	20444	9	US-09-764-891-9422
31	115.8	23.2	8220	10	US-09-797-908-3
32	115.6	23.1	312	9	US-10-102-627-86
33	115.6	23.1	312	9	US-10-102-627-90
34	115.6	23.1	312	10	US-09-764-877-3132
35	115.6	23.1	6145	9	US-10-102-627-87
36	115.6	23.1	6145	9	US-10-102-627-91
37	115.6	23.1	6145	10	US-09-764-877-3133
38	115	23.0	401	9	US-09-946-807-326
39	115	23.0	401	9	US-09-946-807-327
40	115	23.0	401	10	US-09-795-668-326
41	115	23.0	401	10	US-09-795-668-327
42	115	23.0	401	10	US-09-795-686-326
43	115	23.0	401	10	US-09-795-686-327
44	115	23.0	1954	9	US-10-091-572-715
45	115	23.0	1954	9	US-09-764-891-9139

ALIGNMENTS

RESULT 1

US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 15893-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature

Sequence 1075, Ap	
Sequence 1404, Ap	
Sequence 1404, Ap	
Sequence 1404, Ap	
Sequence 7825, Ap	
Sequence 7, Appli	
Sequence 1076, Ap	
Sequence 1077, Ap	
Sequence 1076, Ap	
Sequence 1077, Ap	
Sequence 9422, Ap	
Sequence 3, Appli	
Sequence 86, Appl	
Sequence 90, Appl	
Sequence 87, Appl	
Sequence 91, Appl	
Sequence 3133, Ap	
Sequence 326, App	
Sequence 326, App	
Sequence 327, App	
Sequence 326, App	
Sequence 327, App	
Sequence 715, App	
Sequence 9139, Ap	


```
; LENGTH: 4045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8718

Query Match          24.0%; Score 119.8; DB 9; Length 4045;
Best Local Similarity 73.0%; Pred. No. 9.9e-24;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 133 TTAGAGCTTGCTCTATTCTTCTTCTTCTTCCAAAAAACAACACTACAATTTTGTGTTTGT 192
Db 1603 TTCTACATGGCTTCTTCTCGTTCCTTCTTCTTCCCTTCTTCTTCTTCTCTCTC 1662

Qy 193 TTTGTTTTTTTGTAGACAGAGGCTCTCGAGGTGTACCCAGGCTGGAGTGCAGTGGCGCA 252
Db 1663 TCTCTTTCTTTTGTAGACAGAGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy 253 TTTTCGACTCACCGCAACCTCCGGCTCCGGCTTAAGCGATTCTCCNGCTCAGCCTCCCA 312
Db 1723 TCTCAGCTACTGCAACCTCTGCTCTGAGTTCAAGCGATTCTCTGCTCAGCCTCCCA 1782

Qy 313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
Db 1783 AGTAGCTGGGACTACAAGCATGGCCACCAC 1813

RESULT 8
US-09-764-891-8719
; Sequence 8719, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8719
; LENGTH: 4045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8719

Query Match          24.0%; Score 119.8; DB 9; Length 4045;
Best Local Similarity 73.0%; Pred. No. 9.9e-24;
Matches 154; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 133 TTAGAGCTTGCTCTATTCTTCTTCTTCTTCCAAAAAACAACACTACAATTTTGTGTTTGT 192
Db 1603 TTCTACATGGCTTCTTCTCGTTCCTTCTTCTTCCCTTCTTCTTCTCTCTC 1662

Qy 193 TTTGTTTTTTTGTAGACAGAGGCTCTCGAGGTGTACCCAGGCTGGAGTGCAGTGGCGCA 252
Db 1663 TCTCTTTCTTTTGTAGACAGAGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGCACGA 1722

Qy 253 TTTTCGACTCACCGCAACCTCCGGCTCCGGCTTAAGCGATTCTCTGCTCAGCCTCCCA 312
Db 1723 TCTCAGCTACTGCAACCTCTGCTCTGAGTTCAAGCGATTCTCTGCTCAGCCTCCCA 1782

Qy 313 AGTAGCTGGGACTACAAGCTCGGGACACCAC 343
Db 1783 AGTAGCTGGGACTACAAGCATGGCCACCAC 1813
```

```
RESULT 9
US-08-927-939-78/c
; Sequence 78, Application US/08927939
; Patent No. US20010006640A1
; GENERAL INFORMATION:
; APPLICANT: Grainger, David J.
; APPLICANT: Tatalick, Lauen Marie
```

```
; TITLE OF INVENTION: Compounds and methods to inhibit or
; TITLE OF INVENTION: augment an inflammatory response.
; FILE REFERENCE: 295.022US1
; CURRENT APPLICATION NUMBER: US/08/927,939
; CURRENT FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 78
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27)...(299)
US-08-927-939-78

Query Match          23.9%; Score 119.6; DB 8; Length 1160;
Best Local Similarity 82.5%; Pred. No. 5.6e-24;
Matches 137; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 179 AATTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGGTGTACACCAGGCTGGA 238
Db 767 ATTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTCGCTCTGTCCGCCAGGCTGGA 708

Qy 239 GTGCAGTGGCGGATTTGACTCAGCGCAACCTCCGCTCCGCGCTTAAGCGATTCTCTCT 298
Db 707 GTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCGGTTTCAGCGCAATTTCTCT 648

Qy 299 GCCTCAGCTCCCAAGTAGCTGGGACTACAAGTCCGGACACACG 344
Db 647 GCCTCAGCTCCCGAGTAGCTGGGACTACAAGCGCCGCTCACCG 602
```

```
RESULT 10
US-09-954-531-146/c
; Sequence 146, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 146
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-146
```

```
Query Match          23.9%; Score 119.6; DB 9; Length 1160;
Best Local Similarity 82.5%; Pred. No. 5.6e-24;
Matches 137; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 179 AATTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGTTCTCGAGGTGTACACCAGGCTGGA 238
Db 767 ATTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTCGCTCTGTCCGCCAGGCTGGA 708

Qy 239 GTGCAGTGGCGGATTTGACTCAGCGCAACCTCCGCTCCGCGCTTAAGCGATTCTCTCT 298
Db 707 GTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCGGTTTCAGCGCAATTTCTCT 648
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 28.1345 Seconds
(without alignments)
5450.193 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500

Perfect score: 500

Sequence: 1 taccagtgactgaattcta.....aagctctgagggaactgacgt 500

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	26664	4	US-09-564-805-28
2	120.4	24.1	7680	4	US-09-210-748A-3
3	119.6	23.9	70000	4	US-09-851-896-3
4	115.6	23.1	84495	4	US-09-797-906-3
5	115.4	23.1	8285	4	US-09-732-025-3
6	115.4	23.1	11827	4	US-09-739-455-3
7	115	23.0	38564	4	US-09-734-673-3
8	112.8	22.6	70000	4	US-09-851-896-3
9	111.6	22.3	19011	1	US-08-310-356-36
10	111.6	22.3	19557	5	PCT-US92-06300-1
11	111.4	22.3	20399	4	US-08-938-669A-5
12	111.4	22.3	14636	4	US-09-173-914-6
13	111.2	22.2	246240	2	US-08-724-394A-20
14	111.2	22.2	246240	2	US-08-724-394A-21
15	111.2	22.2	246240	2	US-08-724-394A-22
16	110.8	22.1	5375	3	US-08-757-223-7
17	110.4	22.1	12394	4	US-09-488-856A-10
18	110.4	22.1	84495	4	US-09-797-906-3
19	110.2	22.0	320	1	US-08-629-939-5
20	110.2	22.0	320	1	US-08-759-873-5
21	110	22.0	36159	4	US-09-749-588-3
22	109.8	22.0	246240	2	US-08-724-394A-20
23	109.8	22.0	246240	2	US-08-724-394A-21
24	109.8	22.0	246240	2	US-08-724-394A-22
25	109.4	21.9	568	4	US-09-347-114A-93
26	109.4	21.9	9734	4	US-09-347-114A-80
27	109.2	21.8	14796	4	US-08-975-080-35

c 28	109.2	21.8	14796	4	US-09-630-706-10	Sequence 10, Appl
c 29	109.2	21.8	14796	4	US-09-496-694B-3	Sequence 3, Appl
c 30	109	21.8	602	4	US-09-078-294-27	Sequence 27, Appl
c 31	108.8	21.8	16063	4	US-09-801-052-3	Sequence 3, Appl
c 32	108.6	21.7	36159	4	US-09-749-588-3	Sequence 3, Appl
c 33	108.4	21.7	1624	4	US-08-430-225A-19	Sequence 19, Appl
c 34	108.4	21.7	162450	4	US-09-345-882-1	Sequence 1, Appl
c 35	108.2	21.6	45716	4	US-08-965-048-5	Sequence 5, Appl
c 36	108.2	21.6	45989	4	US-08-965-048-6	Sequence 6, Appl
c 37	108	21.6	16389	4	US-09-741-154-3	Sequence 3, Appl
c 38	108	21.6	18073	4	US-09-078-294-12	Sequence 12, Appl
c 39	108	21.6	50000	4	US-09-146-053-3	Sequence 3, Appl
c 40	108	21.6	112132	4	US-03-741-150-3	Sequence 3, Appl
c 41	107.8	21.6	1001	4	US-09-641-638-629	Sequence 629, App
c 42	107.8	21.6	1001	4	US-09-641-638-630	Sequence 630, App
c 43	107.8	21.6	3742	1	US-08-694-915-5	Sequence 5, Appl
c 44	107.8	21.6	3748	2	US-08-958-240-1	Sequence 1, Appl
c 45	107.8	21.6	20674	4	US-09-641-638-651	Sequence 651, App

ALIGNMENTS

RESULT 1

US-09-564-805-28

; Sequence 28, Application US/09564805

; Patent No. 6333403

; GENERAL INFORMATION:

; APPLICANT: Tavtigan, Sean V.

; APPLICANT: Teng, David H.F.

; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.

; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility

; FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805

; CURRENT FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: US 60/107,468

; PRIOR FILING DATE: 1998-11-06

; PRIOR APPLICATION NUMBER: 09/434,382

; PRIOR FILING DATE: 1999-11-05

; NUMBER OF SEQ ID NOS: 240

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 28

; LENGTH: 26664

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (910)..(13104)

; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:

; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;

; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:

; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:

; NAME/KEY: misc feature

; LOCATION: (13756)..(22917)

; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon

; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:

; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:

; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917

; NAME/KEY: misc feature

; LOCATION: (23045)..(26452)

; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon

; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:

; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation

; OTHER INFORMATION: signal: 26447-26452

; NAME/KEY: variation

; LOCATION: (826)..(23879)

; OTHER INFORMATION: s at positions 826 and 23180 is G or C; Y at

; OTHER INFORMATION: positions 1914, 5568, 7185, 16431, 1857 and 20486

; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at

; OTHER INFORMATION: positions 22211 and 23879 is A or G.

US-09-564-805-28

Query Match 100.0%; Score 500; DB 4; Length 26664;
Best Local Similarity 100.0%; Pred. No. 8.5e-133;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATCAGGTGACGTAATCTATATCTGAAGTAGGAGATCTGTTATTCGTGTTATTACAT 60
Db 1 TATCAGGTGACGTAATCTATATCTGAAGTAGGAGATCTGTTATTCGTGTTATTACAT 60

Qy 61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCAGCAGCTAACAAATGAGCC 120
Db 61 TTTACACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCAGCAGCTAACAAATGAGCC 120

Qy 121 AAGACTCTTGTCTTTAGAGCTTGCTCTTATCTGCTTTTCTTTCCAAAAACACTACAA 180
Db 121 AAGACTCTTGTCTTTAGAGCTTGCTCTTATCTGCTTTTCTTTCCAAAAACACTACAA 180

Qy 181 TTTTGTGTTTGTGTTTGTGTTTGTGACACAGGCTCTGAGGTGTCACCCAGGCTGGAGT 240
Db 181 TTTTGTGTTTGTGTTTGTGTTTGTGACACAGGCTCTGAGGTGTCACCCAGGCTGGAGT 240

Qy 241 GCAGTGGCGGATTTTCGACTCACCCGCAACCTCCGCTCGCGCTTAAGCGATTCCTCTGC 300
Db 241 GCAGTGGCGGATTTTCGACTCACCCGCAACCTCCGCTCGCGCTTAAGCGATTCCTCTGC 300

Qy 301 CTCAGCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACGTAATAAATGATCAAGTT 360
Db 301 CTCAGCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACGTAATAAATGATCAAGTT 360

Qy 361 CTAACATGATGTCATCAAGTAATCAATGAAATAAATAGCAAGCGCTTATGCTAATG 420
Db 361 CTAACATGATGTCATCAAGTAATCAATGAAATAAATAGCAAGCGCTTATGCTAATG 420

Qy 421 CTCATACAAATGATTTCTCTCAATTTAATCTCTCAACACTACAACCACTCTTAACTC 480
Db 421 CTCATACAAATGATTTCTCTCAATTTAATCTCTCAACACTACAACCACTCTTAACTC 480

Qy 481 AAGCTCTGAGGACTGACGT 500
Db 481 AAGCTCTGAGGACTGACGT 500

RESULT 2

US-09-210-748A-3
; Sequence 3, Application US/09210748A
; Patent No. 6335156
; GENERAL INFORMATION:
; APPLICANT: Hermeking, Heiko
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: 14-3-3 SIGMA ARREST THE CELL CYCLE
; FILE REFERENCE: 1107.77810
; CURRENT APPLICATION NUMBER: US/09/210,748A
; CURRENT FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/069,416
; PRIOR FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows-Version 3.0
; SEQ ID NO 3
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-210-748A-3

Query Match 24.1%; Score 120.4; DB 4; Length 7680;
Best Local Similarity 74.3%; Pred. No. 8.1e-25;
Matches 165; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 124 ACTCTGCTTTAGAGCTTCTCTCTATCTGCTTTCTTTCCAAAAACACTACAAATTT 183
Db 5886 AGTCTGATTCCTCAAGACTAGTCTATTCTGTATCTCATAGACAAACAATATATTCAC 5945

Qy 184 TTGTTTGTGTTTGTGTTTGTGAGACAGGCTCTCGAGGTGTACCCAGGCTGGAGTGCA 243
Db 5946 TTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTTGTCTGTACCCAGGCTGGAGTGCA 6005

Qy 244 GTGCGCGGATTTTCGACTCACCGCAACCTCCGCTT-CGCGCTTAAGCGATTCCTCTCCT 302
Db 6006 GTGCGCGCATCTCGGCTCACTGCAACGTCGCTCCGCGGTTCGAGCGATTCCTCTCCT 6065

Qy 303 CAGCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACAGG 344
Db 6066 CAGCTCCCAAGTAGCTGGGACTACAAGCATGTGCCACCATG 6107

RESULT 3

US-09-851-896-3/c
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match 23.9%; Score 119.6; DB 4; Length 70000;
Best Local Similarity 80.5%; Pred. No. 3.3e-24;
Matches 140; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 181 TTTTGTGTTTGTGTTTGTGTTTGTGAGACAGGCTCTCGAGGTGTACCCAGGCTGGAGT 240
Db 31860 TTTTGTGTTTGTGTTTGTGTTTGTGAGACGAGTCTCGCTCTGTCCACCCAGGCTGGAGT 31801

Qy 241 GCAGTGGCGGATTTTCGACTCACCGCAACCTCCGCTCCGCGTTAAGCGATTCCTCTGC 300
Db 31800 GCAGTGGCGGATTTTCGCTCACTGCAACCTCTGCTTCTGGTTCAAGTATTCCTCTGC 31741

Qy 301 CTCAGCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACGTAATAAATGAT 354
Db 31740 CTCAGCTCCCGAGTAGCAGGGATTACAGGAACGTGCGCACCATGCTCAGCTAAT 31687

RESULT 4

US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G

US-09-797-906-3

Query Match 23.1%; Score 115.6; DB 4; Length 84495;
Best Local Similarity 66.7%; Pred. No. 4.9e-23;
Matches 180; Conservative 0; Mismatches 89; Indels 1; Gaps 1;
QY 86 TGAGAGTCAAGTACGAGCTAACAAATGAGCCAGACTCTGCTTTAGAGCTTGTC 145
Db 15760 TGGCAAGCTGACCATGACAGATGGGAGACACAGCTGATGAAGCATTTTTTCCAA 15819
QY 146 TCTATTCTGCTTTCTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTT 205
Db 15820 TTGTGAATATATTATAACAAGAACATTATATCTTTTCTTTTCTTTTCTTTT 15879
QY 206 GAGACAGGCTCTCAGGTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCG 265
Db 15880 GAGACAGAGTCTTGCTCTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCG 15939
QY 266 CAACCTCCGCTCC-GGGCTTAAAGGATTCCTGCTCAGCTCAGCTCCCAAGTAGCTGGAC 324
Db 15940 CAACCTCCGCTCCGCTTCAAGGATTCCTGCTCAGCTCAGCTCCCAAGTAGCTGGAC 15999
QY 325 TACAAGCTCGGACACACCACTGAAAAATGAT 354
Db 16000 TGCAGGTGGCACCACCATGCCAGCTAAT 16029

RESULT 5

US-09-732-025-3
; Sequence 3, Application US/09732025
; Patent No. 6416990
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001011
; CURRENT APPLICATION NUMBER: US/09/732,025
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8285
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(8285)
; OTHER INFORMATION: n = A,T,C or G
US-09-732-025-3

Query Match 23.1%; Score 115.4; DB 4; Length 8285;
Best Local Similarity 74.1%; Pred. No. 2.2e-23;
Matches 146; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 147 CTATTCTGCTTTCTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTT 206
Db 278 CTTTTCCTCTTTTGTGAATCTTTTGAAGTAGTTTGTGTTGTTGTTGTTGTTG 337
QY 207 AGACAGGCTCAGGTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCGC 266
Db 338 AGACAGGCTCAGGTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCGC 397
QY 267 AACCTCCGCTCCGCTTAAAGGATTCCTGCTCAGCTCCCAAGTAGCTGGAGTGA 326
Db 398 AACCTCCGCTCCGCTTAAAGGATTCCTGCTCAGCTCCCAAGTAGCTGGAGTGA 457
QY 327 CAAGCTCGGACACAC 343
Db 458 CAGGCTCATGTCAACAC 474

RESULT 6

QY 127 CTTGCTTTAGAGCTTGCTCTATTCTTTGCTTTTCTTCCAAAAACACTACAATTTT 186

US-09-739-455-3

; Sequence 3, Application US/09739455
; Patent No. 6413756
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000653
; CURRENT APPLICATION NUMBER: US/09/739,455
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 11827
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(11827)
; OTHER INFORMATION: n = A,T,C or G
US-09-739-455-3

Query Match 23.1%; Score 115.4; DB 4; Length 11827;
Best Local Similarity 74.1%; Pred. No. 2.6e-23;
Matches 146; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 147 CTATTCTGCTTTCTTCCAAAAACACTACAATTTTGTGTTTGTGTTTGTGTTT 206
Db 2129 CTTTTCCTCTTTTGTGAATCTTTTGAAGTAGTTTGTGTTGTTGTTGTTGTTG 2188
QY 207 AGACAGGCTCAGGTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCGC 266
Db 2189 AGACAGGCTCAGGTGTCAACAGGCTGGAGTGAGTGGCGGATTCGACTCACCGC 2248
QY 267 AACCTCCGCTCCGCTTAAAGGATTCCTGCTCAGCTCCCAAGTAGCTGGAGTGA 326
Db 2249 AACCTCCGCTCCGCTTAAAGGATTCCTGCTCAGCTCCCAAGTAGCTGGAGTGA 2308
QY 327 CAAGCTCGGACACAC 343
Db 2309 CAGGCTCATGTCAACAC 2325

RESULT 7

US-09-734-673-3
; Sequence 3, Application US/09734673
; Patent No. 6410294
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001020
; CURRENT APPLICATION NUMBER: US/09/734,673
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 38564
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(38564)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-673-3

Query Match 23.0%; Score 115; DB 4; Length 38564;
Best Local Similarity 67.3%; Pred. No. 5.3e-23;
Matches 177; Conservative 0; Mismatches 85; Indels 1; Gaps 1;

QY 127 CTTGCTTTAGAGCTTGCTCTATTCTTTGCTTTTCTTCCAAAAACACTACAATTTT 186


```
FEATURE:
NAME/KEY: intron
LOCATION: 6935..7758
OTHER INFORMATION: /number=5
FEATURE:
NAME/KEY: exon
LOCATION: 7759..7856
OTHER INFORMATION: /number=6
FEATURE:
NAME/KEY: intron
LOCATION: 7857..9443
OTHER INFORMATION: /number=6
FEATURE:
NAME/KEY: exon
LOCATION: 9444..9573
OTHER INFORMATION: /number=7
FEATURE:
NAME/KEY: intron
LOCATION: 9574..10866
OTHER INFORMATION: /number=7
FEATURE:
NAME/KEY: exon
LOCATION: 10867..11081
OTHER INFORMATION: /number=8
FEATURE:
NAME/KEY: intron
LOCATION: 11082..12480
OTHER INFORMATION: /number=8
FEATURE:
NAME/KEY: exon
LOCATION: 12481..12613
OTHER INFORMATION: /number=9
FEATURE:
NAME/KEY: intron
LOCATION: 12614..13701
OTHER INFORMATION: /number=9
FEATURE:
NAME/KEY: exon
LOCATION: 13702..13799
OTHER INFORMATION: /number=10
FEATURE:
NAME/KEY: intron
LOCATION: 13800..14976
OTHER INFORMATION: /number=10
FEATURE:
NAME/KEY: exon
LOCATION: 14977..15115
OTHER INFORMATION: /number=11
FEATURE:
NAME/KEY: intron
LOCATION: 15116..15533
OTHER INFORMATION: /number=11
FEATURE:
NAME/KEY: exon
LOCATION: 15534..15757
OTHER INFORMATION: /number=12
FEATURE:
NAME/KEY: intron
LOCATION: 15758..16949
OTHER INFORMATION: /number=12
FEATURE:
NAME/KEY: exon
LOCATION: 16950..17082
OTHER INFORMATION: /number=13
FEATURE:
NAME/KEY: intron
LOCATION: 17083..17696
OTHER INFORMATION: /number=13
FEATURE:
NAME/KEY: exon
LOCATION: 17697..17764
OTHER INFORMATION: /number=14
```

```
NAME/KEY: intron
LOCATION: 17765..18534
OTHER INFORMATION: /number=14
FEATURE:
NAME/KEY: CDS
LOCATION: join(1776..1854, 2564..2621, 4076..4208,
LOCATION: 6041..6252, 6802..6934, 7759..7856, 9444..9573,
LOCATION: 10867..11081, 12481..12613, 13702..13799,
LOCATION: 14977..15115, 15534..15757, 16950..17082,
LOCATION: 17697..17741)
OTHER INFORMATION: /product="human serum albumin"
FEATURE:
NAME/KEY: exon
LOCATION: 18535..18697
OTHER INFORMATION: /number=15
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 17742..18697
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1737..1775
PUBLICATION INFORMATION:
AUTHORS: Minichetti, P P
AUTHORS: Ruffner, D E
AUTHORS: Kuang, W-J
AUTHORS: Dennison, O E
AUTHORS: Hawkins, J W
AUTHORS: Beattie, W G
AUTHORS: Dugaiczky, A
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN
TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN
TITLE: q11-22 OF CHROMOSOME 4
JOURNAL: J. Biol. Chem.
VOLUME: 261
PAGES: 6747-6757
DATE: 1986
RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011
US-08-310-356-36

Query Match      22.3%; Score 111.6; DB 1; Length 19011;
Best Local Similarity 81.6%; Pred. No. 3.7e-22;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 187 TTTTGTGTTTGTGTTTGTGACACAGGGTCTCGAGGTGTACCCAGGCTGGAGTGCAGTG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3038 TTAATTTTTTTTTTTTTTTTAAACACAGGGTCTCGCTCTGTCGCCAGGCTGGAGTGCAGTG 3097

QY 247 GCGGATTTCGACTCACCAGCAACCTCCGCTCCGGCTTAAGGATTCTCTGCCTCAGC 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3098 GCGCAATCTCGGCTCACTGCAAACTCCGTCGCCGGTTTACGCCATTCTCTGCCTCAGC 3157

QY 307 CTCCCAAGTAGCTGGGACTACAAGCTCGGGACACACAG 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3158 CTCCGAGTAGCTGGGACTACAGCGCCCGCCATCACG 3195

RESULT 10
PCT-US92-06300-1
; Sequence 1, Application PC/TUS9206300
; GENERAL INFORMATION:
; APPLICANT: Hurwitz, David R
; APPLICANT: Nathan, Margret
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: Transgenic Protein Production
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer, Inc.
; STREET: 500 Virginia Ave., Bldg. 3A
; CITY: Ft. Washington
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19034
```

RESULT 11
US-08-938-669A-5
; Sequence 5, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon

PCT-US92-06300-1
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 19002

Qy 307 CTCCCAAGTAGCTGGGACTACAACTCGGACACACG 344
||||| ||||||| ||| ||| |||
Db 3713 CTCCCGAGTAGCTGGGACTACAGGCGCCGCCATCAG 3750

Query Match 22.3%; Score 111.4; DB 4; Length 14636;

GenCore version 5.1.5

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 2075.34 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_21800_22600

Perfect score: 801

Sequence: 1 agtgcctgtcctgtattttt.....agcggaagctttgaccggat 801

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hcg_hum.*

31: em_hcg_inv.*

32: em_hcg_other.*

33: em_hcg_mus.*

34: em_hcg_pln.*

35: em_hcg_rtd.*

36: em_hcg_mam.*

37: em_hcg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	800.6	100.0	118788	9	AC005277	AC005277 Homo sapi
2	728.6	91.0	740	9	AF304371S2	AF304369 Homo sapi
3	434.8	54.3	2793	9	AK094012	AK094012 Homo sapi
4	145.8	18.2	2163	9	AK074244	AK074244 Homo sapi
5	145.8	18.2	2734	9	AK094333	AK094333 Homo sapi
6	145.8	18.2	2893	9	AF308694	AF308694 Gorilla g
7	145.8	18.2	2908	9	AF304370	AF304370 Homo sapi
8	145.8	18.2	2908	9	AF308698	AF308698 Pan trogl
9	145.8	18.2	2976	9	AK001392	AK001392 Homo sapi
10	145.8	18.2	2992	6	AX405825	AX405825 Sequence
11	145.8	18.2	2997	9	BC001939	BC001939 Homo sapi
12	145.8	18.2	3006	9	BC004158	BC004158 Homo sapi
13	144.2	18.0	2907	9	AK094687	AK094687 Homo sapi
14	139	17.4	167336	2	AC128361	AC128361 Rattus no
15	138.6	17.3	34593	10	AF348157	AF348157 Mus muscu
16	138.6	17.3	196566	10	AL663045	AL663045 Mouse DNA
17	137	17.1	188836	2	AC097911	AC097911 Rattus no
18	102.2	12.8	2712	10	AF308696	AF308696 Mus muscu
19	46	5.7	125020	9	AF429315	AF429315 Homo sapi
20	45.4	5.7	144048	9	AC013762	AC013762 Homo sapi
21	45	5.6	173016	2	AC019123	AC019123 Homo sapi
22	44.2	5.5	214025	9	AC026782	AC026782 Homo sapi
23	43.8	5.5	120703	9	AC002364	AC002364 Homo sapi
24	43.8	5.5	164846	2	AC023007	AC023007 Homo sapi
25	43.8	5.5	248366	2	AL805901	AL805901 Mus muscu
26	43.4	5.4	179838	2	AC130575	AC130575 Rattus no
27	43	5.4	199847	2	AC127320	AC127320 Mus muscu
28	42.8	5.3	157173	9	AC009501	AC009501 Homo sapi
29	42.6	5.3	182411	9	AC068044	AC068044 Homo sapi
30	42.4	5.3	119191	9	AP004782	AP004782 Homo sapi
31	42.2	5.3	120512	9	AL359852	AL359852 Human DNA
32	42	5.2	183540	2	AC026114	AC026114 Homo sapi
33	42	5.2	188972	9	AC020611	AC020611 Homo sapi
34	41.8	5.2	67970	3	PFMAL1P3	AL031746 Plasmodiu
35	41.8	5.2	138915	2	AF005246	AF005246 Oryza sat
36	41.8	5.2	167764	2	AF004303	AF004303 Oryza sat
37	41.8	5.2	236120	14	AF063866	AF063866 Melanoplu
38	41.6	5.2	161355	2	AC011154	AC011154 Homo sapi
39	41.6	5.2	182374	2	AP003968	AP003968 Homo sapi
40	41	5.1	140106	9	AC096713	AC096713 Homo sapi
41	41	5.1	209060	2	AC019005	AC019005 Homo sapi
42	41	5.1	224448	2	PFMAL4P4	AL035477 Plasmodiu
43	40.8	5.1	206915	10	AL663081	AL663081 Mouse DNA
44	40.8	5.1	210185	2	AC125321	AC125321 Mus muscu
45	40.6	5.1	6370	6	AX344345	AX344345 Sequence

ALIGNMENTS

RESULT 1
AC005277
LOCUS AC005277 118788 bp DNA linear PRI 23-JUL-1998
DEFINITION Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 118788)
AUTHORS Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK.597_M_12
JOURNAL Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraigery,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Melidram,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo.A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 118788)
 Birren,B., Fasman,K., Linton,L., Nusbbaum,C., Lander,E., Allen,N.,
 Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
 Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
 FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
 Gensheimer,S., Geraigery,K., Gilmartin,T., Grant,G., Hagos,B.,
 Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Melidram,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
 Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
 Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
 Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
 Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
 Vo.A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
 Zody,M.

TITLE
JOURNAL

Direct Submission
 Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 23, 1998 this sequence version replaced gi:335015.
 All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

FEATURES
source

Location/Qualifiers
 1. .118788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="hRPK.597 M.12"
 /clone.lib="RPCI-II human BAC library"

795..884

/rpt_family="MIR"

complement(932..1000)

/rpt_family="MIR"

2044..2076

/rpt_family="AT_rich"

2237..2713

/rpt_family="MIR34"

2718..3477

/rpt_family="12"

complement(3478..3821)

/rpt_family="MLT1A1"

3822..4007

/rpt_family="12"

complement(4015..4281)

/rpt_family="MER89"

complement(4282..4321)

/rpt_family="(TAAA)n"

complement(4328..4608)

repeat_region
 complement(4609..4895)
 /rpt_family="MER89"
 repeat_region
 complement(4896..5186)
 /rpt_family="AluSp"
 5219..5479
 /rpt_family="L2"
 repeat_region
 complement(5481..5740)
 /rpt_family="MER34"
 repeat_region
 complement(5963..6275)
 /rpt_family="AluSp"
 repeat_region
 complement(6411..6494)
 /rpt_family="MER5A"
 6830..7013
 /rpt_family="MLT1C"
 7014..7318
 /rpt_family="AluYb8"
 7319..7581
 /rpt_family="MLT1C"
 7721..8032
 /rpt_family="AluJb"
 8039..8333
 /rpt_family="AluSq"
 9431..9460
 /rpt_family="AT_rich"
 9662..9829
 /rpt_family="MIR"
 complement(11302..11728)
 /rpt_family="MER39"
 complement(11824..11884)
 /rpt_family="LTR16C"
 complement(11928..12108)
 /rpt_family="LTR16C"
 12453..12725
 /rpt_family="AluJb"
 13377..13785
 /rpt_family="MSTB"
 14118..14153
 /rpt_family="(CA)n"
 complement(15320..15620)
 /rpt_family="AluJo"
 15683..15759
 /rpt_family="(CA)n"
 complement(15838..15893)
 /rpt_family="MER91A"
 complement(15915..16080)
 /rpt_family="MER91A"
 complement(16474..16544)
 /rpt_family="L2"
 17907..17920
 /rpt_family="L1ME"
 18271..18311
 /rpt_family="L2"
 complement(18368..18529)
 /rpt_family="AluJb"
 complement(18530..18560)
 /rpt_family="(CA)n"
 complement(18561..18681)
 /rpt_family="AluJb"
 18760..19065
 /rpt_family="AluSq"
 complement(19638..19896)
 /rpt_family="MIR"
 complement(19971..20260)
 /rpt_family="L1ME3A"
 20263..20552
 /rpt_family="MER58B"
 complement(20553..20816)
 /rpt_family="L1ME3A"
 complement(20821..21069)
 /rpt_family="L1M4"
 complement(21071..21100)
 /rpt_family="(TAAA)n"

```

repeat_region complement(21101..21338)
repeat_region /rpt_family="AluSx"
repeat_region complement(21348..21598)
repeat_region /rpt_family="L1M4"
repeat_region 21799..22104
repeat_region /rpt_family="AluJo"
repeat_region 22151..22562
repeat_region /rpt_family="MSTA"
repeat_region 23978..24287
repeat_region /rpt_family="AluSq"
repeat_region 24852..25199
repeat_region /rpt_family="L2"
repeat_region complement(25697..25985)
repeat_region /rpt_family="MLTID"
unsure 25816..25895
repeat_region /note="Single-stranded coverage."
repeat_region complement(25994..26099)
repeat_region /rpt_family="purine-rich"
repeat_region complement(26122..26291)
repeat_region /rpt_family="MLTID"
repeat_region 26312..26337
repeat_region /rpt_family="n" (CAAAA)n"
repeat_region 26456..26480
repeat_region /rpt_family="n" (GGGAA)n"
repeat_region complement(26576..26781)
repeat_region /rpt_family="MER6B"
repeat_region complement(26912..26990)
repeat_region /rpt_family="MER21B"
repeat_region 26991..27290
repeat_region /rpt_family="AluSx"
repeat_region complement(27291..27975)
repeat_region /rpt_family="MER21B"
repeat_region 28034..28197
repeat_region /rpt_family="LTR28"
repeat_region 28266..28795
repeat_region /rpt_family="LTR28"
repeat_region 28843..28916
repeat_region /rpt_family="LTR28"
repeat_region 28920..29005
repeat_region /rpt_family="MIR"
repeat_region 29140..29233

Query Match 100.0%; Score 800.6; DB 9; Length 118788;
Best Local Similarity 99.9%; Pred. No. 1.3e-196;
Matches 800; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGCCGTGCTGATTTTCAACAGAGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 60
Db 99452 AGTGCCGTGCTGATTTTCAACAGAGCTGTGGCCACAGTCAATCTGCATGGTCAGAT 99511

QY 61 TCATTGTTAGGACTAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTGTATGC 120
Db 99512 TCATTGTTAGGACTAATGCTTTAAGCCTCCTATAAACTTTTTTTTTTTTTTTTGTATGC 99571

QY 121 CGAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCCATGATACTCTTTTGCTA 180
Db 99572 CGAGCCTTTGTGTAAGTCTACTTGAAGGGTTTCAGGGTTCCATGATACTCTTTTGCTA 99631

QY 181 TAAAGAGGATGACATGTAAATACCTTTATGGTTAAATTAATTTGGCTTTATATTAG 240
Db 99632 TAAAGAGGATGACATGTAAATACCTTTATGGTTAAATTAATTTGGCTTTATATTAG 99691

QY 241 CTCCTCAAAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTTCTTGCTCTGTCCTCAA 300
Db 99692 CTCCTCAAAAGCAAGCAGGAGACAGAAATTTCTGAGTTGCTTCTTGCTCTGTCCTCAA 99751

QY 301 AGCAGACATAGCCTCTGAACCATCAGCAGTCTTCTAGTGGCAGTGACTCTTCTCTCT 360
Db 99752 AGCAGACATAGCCTCTGAACCATCAGCAGTCTTCTAGTGGCAGTGACTCTTCTCTCT 99811

QY 361 TCTCTTCTGAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGACACRTTTGGGCA 420
Db 99812 TCTCTTCTGAGCCCGACACGCTCTGCTACTGGACTGTGGTGGGACACRTTTGGGCA 99871

```

```

QY 421 GCTGTGCCGTCAATTACGAGAGACCAGGTGGACAGGGTCTCTGGGCACCCCTGGCTGTGTGT 480
Db 99872 GCTGTGCCGTCAATTACGAGAGACCAGGTGGACAGGGTCTCTGGGCACCCCTGGCTGTGTGT 99931

QY 481 TGTGTCCCACTGTCACGACGATCACACAGGTGAGTGTGGCTGGACCAACAGCTGG 540
Db 99932 TGTGTCCCACTGTCACGACGATCACACAGGTGAGTGTGGCTGGACCAACAGCTGG 99991

QY 541 AGCCTGGAGGAGGCACTGCCAGTTGAGTTGGCCCTTTGGCTGGCTCTTTCTCCGCTT 600
Db 99992 AGCCTGGAGGAGGCACTGCCAGTTGAGTTGGCCCTTTGGCTGGCTCTTTCTCCGCTT 100051

QY 601 CCAAACTTGCACAGAGCTTTTGTATCTCATCTCTGGCTAGGAAATGGTTTTTGGAAAAAC 660
Db 100052 CCAAACTTGCACAGAGCTTTTGTATCTCATCTCTGGCTAGGAAATGGTTTTTGGAAAAAC 100111

QY 661 TCAACATAGTCTTCTGCGCCACAGAAATGCTCTCTCTCTGTTTCTGTTTCTTCTCTG 720
Db 100112 TCAACATAGTCTTCTGCGCCACAGAAATGCTCTCTCTCTGTTTCTTCTGTTTCTTCTCTG 100171

QY 721 AGCAGCAGAGTTTGTAGTTTACCCAGCCTTCTTCTGAGTCTTGAATCTCACACGGCTGTCT 780
Db 100172 AGCAGCAGAGTTTGTAGTTTACCCAGCCTTCTTCTGAGTCTTGAATCTCACACGGCTGTCT 100231

QY 781 CAGCGAAGCTTTGACCGGAT 801
Db 100232 CAGCGAAGCTTTGACCGGAT 100252

RESULT 2
AF304371S2 740 bp DNA linear PRI 19-OCT-2000
LOCUS Homo sapiens putative prostate cancer susceptibility protein
DEFINITION HPC2/BLAC2 gene, exon 17 and partial cds.
ACCESSION AF304369
VERSION AF304369.1 GI:10880929
KEYWORDS 2 of 2
SEGMENT Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
Gupte,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
A strong candidate prostate cancer susceptibility gene at
chromosome 17p
Unpublished
2 (bases 1 to 740)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
Gupte,J.S., Hu,R., Iliev,D., Janacki,T., Kort,E.N., Laity,K.E.,
Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
and Cannon-Albright,L.A.
Direct Submission
Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt
Lake City, UT 84108, USA
Lake City, UT 84108, USA
Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

```


Qy	458	CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCACGCAGATCACACACGG	512
Db	845	CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCACGCAGATCACACACGG	899
RESULT 5			
AK094333			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 156;			
Conservative			
1;			
Mismatches			
18;			
Indels			
0;			
Gaps			
0;			
Qy	338	AGTGGCAGTGACTCTCTTCCTCTTCTCTTCCTGCAGCCCCGACACGCTCTCTGCTACTGGAC	397
Db	1314	ATTGCAAAATGTCACTGAGTGCACACTTGTCAAATAAGCCCCGACACGCTCTCTGCTACTGGAC	1373
Qy	398	TGTGGTGAGGCACRTTTGGGCAGCTGTGCGGTCTATTACGGAGACCAGGTGGACAGGTC	457
Db	1374	TGTGGTGAGGCACATTTGGGCAGCTGTGCGGTCTATTACGGAGACCAGGTGGACAGGTC	1433
Qy	458	CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCACGCAGATCACACACGG	512
Db	1434	CTGGCACCCTGGCTGCTGTGTTGTCTCCACCTGCACGCAGATCACACACGG	1488
RESULT 6			

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk
Email: cgepbs-r@mail.nih.gov

Tissue Procurement: DCTP/DTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywninski, Reta Kutische, Oliver Lee, Soo Len Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Fawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasya van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 9 Row: 1 Column: 17
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 7022621.

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.	1. High Accuracy: The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.
2. Scalability: The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.	2. Scalability: The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.
3. Interpretability: The model's decisions are transparent and explainable, allowing users to understand the underlying logic and reasoning.	3. Interpretability: The model's decisions are transparent and explainable, allowing users to understand the underlying logic and reasoning.
4. Robustness: The model is highly resistant to adversarial attacks and data poisoning, ensuring reliable performance in real-world scenarios.	4. Robustness: The model is highly resistant to adversarial attacks and data poisoning, ensuring reliable performance in real-world scenarios.
5. Efficiency: The model is optimized for fast inference and low resource consumption, making it ideal for deployment on edge devices.	5. Efficiency: The model is optimized for fast inference and low resource consumption, making it ideal for deployment on edge devices.

```

Location/Qualifiers
1. .2997
  /organism="Homo sapiens"
  /db_xref="LocusID:60528"
  /db_xref="taxon:9606"
  /clone="WGC:4102 IMAGE:2820640"
  /tissue_types="Lung, small cell carcinoma"
  /clone_lib="NIH MGC 7"
  /lab_host="DH10B-R"
  /note="Vector: pOTB7"
61. .2541

```

CDS

```

/codon_start=1
/products "putative cancer cancer susceptibility protein"
/protein_id="AAH01939.1"
/db_xref="GI:12804973"
/translation="MWAQGLSLRSAGRTMSGQRTISQAAPRRPRKDLRLHRLTR
KRGSPSGGPNTVLQVVAAGSDGSAALVYSEFNRLPNCGEVYQVFGSHKVLK
ARLDNLTFLTRMHSNVGGLSGMLTDLKTEGLPKCVLSGPOLKYLEAIKIQSLKLG
ELAVRPHANPEYEDETMTVYQIPIHSQRQKQKQWQSPERPLRSLRSPRSSDELN
ENPDPHGVSQRGVRDSSLVVATIKLHLKRGNVLFLKAKEMGLPVTGAALPIA
AXDKGSIITHEGRSILAEELSTPPGAAALFVVCPQESFTQPTCENATFQYQKAD
APVALVWNPAPSLVDVRYQWQMERFGDTQLHVLNENCAVHLSRHLIQTLNLI
HPDIPLPALTFRCKBEGPSTLVPMVQGLCKLYQLRPREMDRAITCNPEFIVEA
DQLSSQSVQVEYRSGAODGAPAEKRSQYPEIPLGTGSAIPMKINVSATLVNPSI
LTLNLLDCGEFTGFCGLCRHVGQDVRGLGTLANVPVSHLADHTGLPSILLQERAL
ASLGLPLHLLVVAQNOLKAMQVHQNOCEVLLHISIMPAKLCQEGAEISSPAVREL
ISLLEATDLDEEFTCLVRHCKHAGCAALVTSYDGMVVKVYSGDTPMCSALVRMGKDATI
LHUNTLEDQEGEAEVKEHTSHSTQAI SVGRMMAEATIMLNHFQRYAKVPLSPNFS
EKGVGAFDMFGQVCFDGFPTMPLKIPLKALFAGDITEEERREKRLRQVRAALLRS
LAGLEDGFPQMKVQDGFPTQAKKVQA"
a      806 c      853 g      598 t

```

BASE COUNT
ORIGIN

	Query Match	18.2%	Score 145.8	DB 9	Length 2997	
	Best Local Similarity	89.1%	Pred. No. 6.4e-27			
	Matches 156	Conservative 1	Mismatches 18	Indels 0	Gaps 0	
Qy	338	AGTGGCAGTGACTCTCTTCCTCTCTCTTCTGCGAGCCCGACACGCTCTCTGCTACTGGAC	397			
Db	1546	ATTGGAATGTCAGTGCCACACTTGTCAACATATAGCCCCGACACGCTCTGCTACTGGAC	1605			
Qy	398	TGTGGTGAGGGCAGCTTTGGGCAGCGTGTGCCGTCAATTACGGAGACCAAGGTGGACAGGGTC	457			
Db	1606	TGTGGTGAGGGCAGCTTTGGGCAGCGTGTGCCGTCAATTACGGAGACCAAGGTGGACAGGGTC	1665			
Qy	458	CTGGGCACCCCTGGCTGTGTTTGTGTGCCCTCTGACGCAGATCACCCACCGG	512			

Db	1666	CTGGGCACCTGGCTGCTGTGTGTGTGCTCCACCTGCACGAGATCACACAGG	1720
RESULT 12	BC004158		
LOCUS	BC004158	3006 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, putative prostate cancer susceptibility protein, clone MGC:2441 IMAGE:2820640, mRNA, complete cds.		
ACCESSION	BC004158		
VERSION	BC004158.1	GI:13278770	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3006)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
	Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NTH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapsb-femail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadansystemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 2 Row: c Column: 13 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7022621.		

FEATURES

```

1. .3006
/organism="Homo sapiens"
/db_xref="LocusID:60528"
/db_xref="taxon:9606"
/clone="MGC:2441 IMAGE:2820640"
/tissue_type="lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
70. .2550
/codon_start=1
/product="putative prostate cancer susceptibility protein"

```

CD5

```

/codon_start=1
/product="putative prostate cancer susceptibility protein"
/protein_id="AA04158.1"
/db_xref="GI:13278771"
/translation="MWAQLSLRRSAAGRTSSQGRITSSQAPARPRKDPRLRHRLTRTE
KRGPSGQSGGVTVLVVAAGSRDGAALTVSEFNRYPLNCGCEQVQRMLQSHKILV
AKLDNI FLTRMWSNVGGLSGMILTKETGLPKCVSLQPLQKLEAKIFGSGLKV
IELAVRPAASEYEDETDTSVVAIFPIHSEQRGKHQWQSPERLSRSPVSGSDSELN
INEMHLPHGVSQVSDRVSQVLAIFIKLHLKRGNFVLVLAKEMGLPVGTAAALPIA
ADWDGKSLTHRETRTELABELCTPPDGAALVWVECDPESTIQICENATRTQVQKAD
APVALVHMAAPASVLVSRYOOWMERGPTDTHVLNENESSVHLNRSHKIQTOLNLI
HPIIFPLTSLFRCKEKGDTSLVPMQCEGLKCYQLRERMQARDITCNPERFIVEA
LOLNPQSVQSVYRSPADGCPAPAKESQYPELIFGTGSAIPMKIRNYSATLVNLSL
DTSLLLDGEGTFGQLCRHYGQDQDVRLGTLAAVFVSHLHADHHTGLPSLLQROERAL
ASGLGPHPLVAFNAPQLWQIQVQNCQCEGLVHLSMIYPAKCGQEAELSSFAVERL
ISLTLQCDLBEQFTCLVRKCHAFGACGLVHTSGSWKVAEVDGTMPCBAELVSGKQDNL
LKEVHTLBDGUEBUEDEKTHSTTSQALSVGRWMAEAPIMLUNHESQYAKVLPSPFNS
LIGFATVPMVKVCFDPEETMPLTLPPLKALFAGDIEMERREKRELQVRAALLSRE
LAGGLEDGEPQQKQAHTEPEQAKVRAQ"
a      808 c 858 q 598 t

```

	BASE COUNT	ORIGIN
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000
11	0.000	0.000
12	0.000	0.000
13	0.000	0.000
14	0.000	0.000
15	0.000	0.000
16	0.000	0.000
17	0.000	0.000
18	0.000	0.000
19	0.000	0.000
20	0.000	0.000
21	0.000	0.000
22	0.000	0.000
23	0.000	0.000
24	0.000	0.000
25	0.000	0.000
26	0.000	0.000
27	0.000	0.000
28	0.000	0.000
29	0.000	0.000
30	0.000	0.000
31	0.000	0.000
32	0.000	0.000
33	0.000	0.000
34	0.000	0.000
35	0.000	0.000
36	0.000	0.000
37	0.000	0.000
38	0.000	0.000
39	0.000	0.000
40	0.000	0.000
41	0.000	0.000
42	0.000	0.000
43	0.000	0.000
44	0.000	0.000
45	0.000	0.000
46	0.000	0.000
47	0.000	0.000
48	0.000	0.000
49	0.000	0.000
50	0.000	0.000
51	0.000	0.000
52	0.000	0.000
53	0.000	0.000
54	0.000	0.000
55	0.000	0.000
56	0.000	0.000
57	0.000	0.000
58	0.000	0.000
59	0.000	0.000
60	0.000	0.000
61	0.000	0.000
62	0.000	0.000
63	0.000	0.000
64	0.000	0.000
65	0.000	0.000
66	0.000	0.000
67	0.000	0.000
68	0.000	0.000
69	0.000	0.000
70	0.000	0.000
71	0.000	0.000
72	0.000	0.000
73	0.000	0.000
74	0.000	0.000
75	0.000	0.000
76	0.000	0.000
77	0.000	0.000
78	0.000	0.000
79	0.000	0.000
80	0.000	0.000
81	0.000	0.000
82	0.000	0.000
83	0.000	0.000
84	0.000	0.000
85	0.000	0.000
86	0.000	0.000
87	0.000	0.000
88	0.000	0.000
89	0.000	0.000
90	0.000	0.000
91	0.000	0.000
92	0.000	0.000
93	0.000	0.000
94	0.000	0.000
95	0.000	0.000
96	0.000	0.000
97	0.000	0.000
98	0.000	0.000
99	0.000	0.000
100	0.000	0.000

Query Match	Score	DB	Length
18.2%	145.8	9	3006


```
Db 127052 TTCCTTCTCTGCTTTTGATGCCAGCATGTTGTTGATGCATATCCTTAAAGGCTTCAGG 126933
Qy 158 GTTCCATGATGACTCTTTTGCTATTAAGAGGATGACACATGTAAATACACCTTTATGGTT 217
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126992 GTTCTTTGATTTTCCCTTAAACAGAGAAGACAGACTGCTCTTAAACGACTCATGTGGTT 126933
Qy 218 AAATTAATGGCTTTTATATTAGCTCTCTCAAGCAAGCAGGAGAGACAGAAATTTCTGC 277
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126932 AAG-TAGTTGACTTCATATGCACTCCACAAAGATGGTGGGTGT---AGCTGTGGAA 126877
Qy 278 AGTTGCTTCTTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 337
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126876 GATGACGTGGTGGCCACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126817
Qy 338 AGTGGCAGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 397
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126816 GGAAGTCTAAAGCCCTGTTGTGACTGCTCTCTGACGCCCTGACAAGTCTGTGCTGTGAT 126757
Qy 398 TGTGTGTAGGGCAGCTTTGGGCGAGTGTGCCCTCATTTAGGGAGACCGGTGCACAGGCTC 457
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126756 TGTGAGAGAGGCACCTTTTGGCGAGTTATGCCGTCTATTGGACAGCAATAGACAGATC 126697
Qy 458 CTGGCAGCCCTGGCTGCTGTTGTTGTGCTCCACCTGCAGCAGATCACCACACGGTGTAGT 517
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126696 CTCTGCACTTACTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 126637
Qy 518 GTTGGGC 524
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126636 GTGAGAC 126630

RESULT 15
AF348157
LOCUS
DEFINITION
Mus musculus putative prostate cancer susceptibility protein
(Elac2) gene, complete cds, alternatively spliced.
ACCESSION
AF348157
VERSION
AF348157.1 GI:13540341
KEYWORDS
Mus musculus.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (sites)
Frank,D.C., Swedlund,B., Dumont,M., Tavtigian,S.V., Simard,J.,
Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R.,
Chen,Y., Dayananth,P., Desrochers,M., Farnham,J.M., Frye,C.,
Gaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swensen,J., Thomas,A.,
Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Mouse Elac2-containing genomic DNA
unpublished
JOURNAL
2 (bases 1 to 34593)
REFERENCE
Frank,D.C., Swedlund,B., Dumont,M., Tavtigian,S.V., Simard,J.,
Teng,D.H.F., Baumgard,M., Beck,A., Camp,N.J., Carillo,A.R.,
Chen,Y., Dayananth,P., Desrochers,M., Farnham,J.M., Frye,C.,
Gaffari,S., Gupta,J.S., Hu,R., Iliev,D., Janecki,T., Kort,E.N.,
Laity,K.E., Leavitt,A., Leblanc,G., McArthur-Morrison,J.,
Pederson,A., Penn,B., Peterson,K.T., Reid,J.E., Richards,S.,
Schroeder,M., Smith,R., Snyder,S.C., Swensen,J., Thomas,A.,
Tranchant,M., Woodland,A.-M., Labrie,F., Skolnick,M.H.,
Neuhausen,S., Rommens,J. and Cannon-Albright,L.A.
Direct Submission
JOURNAL
Submitted (09-FEB-2001) Myriad Genetics, Inc., 320 Wakara Way, Salt
Lake City, UT 84108, USA
FEATURES
Location/Qualifiers
1..34593
/organism="Mus musculus"
/db_xref="taxon:10090"
<8470..>31277
gene
```

```
/gene="Elac2"
join(<8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30624,
31242..>31277)
/gene="Elac2"
/product="putative prostate cancer susceptibility protein"
/note="alternatively spliced 25 exon form"
join(8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30624,
31242..31277)
/gene="Elac2"
/note="25 exon form"
/codon_start=1
/product="putative prostate cancer susceptibility protein"
/protein_id="AAK29420.1"
/db_xref="GI:13540342"
/translation="MWALRSLLRPLGLRTWMSQGSARRPPPKDPLRLHRTREKRGPG
GGPNTVYLVVAAAGRDAGALYVSEYNYLFCNGEGVORLMOEHLKVARLDNI FL
TRHWSNVGGLCGMILITKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGLAVRPH
SAPYKDETMTVYQVPIHSERRCQKQSPSPKRLSPKSSQSGSANGQCQPE
DSSAGANRKAWGRDPSLVAVCKHLRKNFLVKAKELGLPVGTAAIPIAIAVKD
GKSITYEGREIAAEELCTPPDPGLVFIWVCEPDGFFILPCENDTFKRYQAEADPA
LVVHIAPESVLIDRSYQWMPFGDPQTHLILNENCPNVHLRSHKIQTQSLIHPDI
PQLTFSVSKSEEGSTLSVPTVRGECILKYQLRPKEWQDITLDCNTDEFTAEALELP
SFQESVEEYRKVNQENPAPAKRSQYFEIVFLGTGSAIPMKIRNVSTLVNLSPKSV
LLDCGEGTFOQLCRHYGQCDRLVCLTAVFVSHLHADHTGLNLLQREHALASLG
KPFQPLLVAPTQLRAWLQYHNHCQIILHVSIMIPAKCLOKGAESVNTLIERLISLL
LETCDLEEFQTCVLRHCKHAFGCVLHSSGKVVYSGDTMPCFALVQMGKATLIIHE
ATLEDGLEEAEVKTHTTQAINVGMNMAEFIMLHFSQYAKIPLFSPDFNEKVG
IAFDHMKVCFGDEFTVPEKLIPLKALFAGDIEENVERRERELRVRAALLTQQADSP
EDREPQOKRAHTIDEPHSPQSKKESVANTLCARV"
join(<8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30639)
/gene="Elac2"
/product="putative prostate cancer susceptibility protein"
/note="alternatively spliced 24 exon form"
join(8470..8690,9077..9127,9266..9336,9948..10012,
10957..11014,12096..12164,12933..13064,14902..14960,
16609..16667,16746..16818,17370..17482,18671..18766,
21539..21677,21757..21842,23553..23671,24619..24715,
27291..27429,27763..27801,27930..28039,28553..28652,
28733..28853,29101..29179,29710..29854,30406..30639)
/gene="Elac2"
/product="putative prostate cancer susceptibility protein"
/note="24 exon form"
/codon_start=1
/product="putative prostate cancer susceptibility protein"
/protein_id="AAK29421.1"
/db_xref="GI:13540343"
/translation="MWALRSLLRPLGLRTWMSQGSARRPPPKDPLRLHRTREKRGPG
GGPNTVYLVVAAAGRDAGALYVSEYNYLFCNGEGVORLMOEHLKVARLDNI FL
TRHWSNVGGLCGMILITKETGLPKCVLSGPPQLEKYLEAIKIFSGPLKGLAVRPH
SAPYKDETMTVYQVPIHSERRCQKQSPSPKRLSPKSSQSGSANGQCQPE
DSSAGANRKAWGRDPSLVAVCKHLRKNFLVKAKELGLPVGTAAIPIAIAVKD
GKSITYEGREIAAEELCTPPDPGLVFIWVCEPDGFFILPCENDTFKRYQAEADPA
LVVHIAPESVLIDRSYQWMPFGDPQTHLILNENCPNVHLRSHKIQTQSLIHPDI
PQLTFSVSKSEEGSTLSVPTVRGECILKYQLRPKEWQDITLDCNTDEFTAEALELP
SFQESVEEYRKVNQENPAPAKRSQYFEIVFLGTGSAIPMKIRNVSTLVNLSPKSV
LLDCGEGTFOQLCRHYGQCDRLVCLTAVFVSHLHADHTGLNLLQREHALASLG
KPFQPLLVAPTQLRAWLQYHNHCQIILHVSIMIPAKCLOKGAESVNTLIERLISLL
LETCDLEEFQTCVLRHCKHAFGCVLHSSGKVVYSGDTMPCFALVQMGKATLIIHE
ATLEDGLEEAEVKTHTTQAINVGMNMAEFIMLHFSQYAKIPLFSPDFNEKVG
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 1295.47 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-28_COPY_1_500
Perfect score: 500
Sequence: 1 taccagtgactgaattctta.....aagctctgaggagactgacgt 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEnbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	118788	9	AC005277	AC005277 Homo sapi
2	131	26.2	157159	2	AC026279	AC026279 Homo sapi
3	131	26.2	157590	9	AC087525	AC087525 Homo sapi
4	131	26.2	171418	2	AC023591	AC023591 Homo sapi
5	128.6	25.7	180707	2	AC017038	AC017038 Homo sapi
6	128.6	25.7	184338	9	AC060764	AC060764 Homo sapi
7	126.4	25.3	110000	2	AC009771.2	Continuation (3 of
8	126.4	25.3	110000	2	AC012386_0	Continuation (3 of
9	126.2	25.2	148068	9	CNS01RHN_0	AC012386 Homo sapi
10	126.2	25.2	179937	9	CNS05TCV	AL161871 Human chr
11	125.8	25.2	115396	9	AC010366	AL355840 Human chr
12	125.8	25.2	214139	2	AC022129	AC022129 Homo sapi
13	125.6	25.1	74089	9	AL589762	AL589762 Human DNA
14	125.6	25.1	208309	9	CNS01DSC	AL121767 Human chr
15	125.4	25.1	153185	2	AC073228	AC073228 Homo sapi
16	125.4	25.1	178653	2	AC090088	AC090088 Homo sapi
17	125.4	25.1	178980	9	AP001591	AP001591 Homo sapi
18	125.4	25.1	184057	9	AP000753	AP000753 Homo sapi
19	125.4	25.1	184059	2	AC090382	AC090382 Homo sapi
20	125	25.0	65612	9	AL136304	AL136304 Human DNA
21	124.8	25.0	199882	9	AL354720	AL354720 Human DNA
22	124.6	24.9	191468	9	AC092631	AC092631 Homo sapi
23	124.2	24.9	171779	2	AC016000	AC016000 Homo sapi
24	124.2	24.8	50721	2	AC069106	AC069106 Homo sapi
25	124.2	24.8	96512	9	AC022539	AC022539 Homo sapi
26	124.2	24.8	154359	2	AC021718	AC021718 Homo sapi
27	124.2	24.8	158568	9	AC069541	AC069541 Homo sapi
28	124.2	24.8	159895	9	AC091772	AC091772 Homo sapi
29	124.2	24.8	189786	2	AC068015	AC068015 Homo sapi
30	124	24.8	175642	2	AC116934	AC116934 Papio cyn
31	124	24.8	181532	9	AL499582	AL499582 Human DNA
32	123.6	24.7	59915	2	AC111187	AC111187 Homo sapi
33	123.4	24.7	146957	9	AC084264	AC084264 Homo sapi
34	123.4	24.7	186091	9	AC011242	AC011242 Homo sapi
35	123.2	24.6	149679	2	AC080115	AC080115 Homo sapi
36	123.2	24.6	155276	2	AC015576	AC015576 Homo sapi
37	123.2	24.6	161149	2	AC012028	AC012028 Homo sapi
38	123.2	24.6	165434	9	AC022448	AC022448 Homo sapi
39	123.2	24.6	178516	9	AC109993	AC109993 Homo sapi
40	123.2	24.6	190740	2	HSAS24G17	AL454722 Homo sapi
41	123	24.6	164084	9	AL354953	AL354953 Human DNA
42	123	24.6	173853	2	AC018999	AC018999 Homo sapi
43	123	24.6	193803	9	AC022819	AC022819 Homo sapi
44	122.8	24.6	178530	9	AC018738	AC018738 Homo sapi
45	122.8	24.6	188025	2	AC074185	AC074185 Homo sapi

ALIGNMENTS

RESULT 1 AC005277 118788 bp DNA linear PRI 23-JUL-1998
AC005277 Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.
LOCUS AC005277
DEFINITION Homo sapiens chromosome 17, clone hRPK.597_M_12, complete sequence.
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 118788)
REFERENCE
AUTHORS Birren,B., Fasmann,K., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK.597_M_12
JOURNAL Unpublished

```

REFERENCE
AUTHORS
2 (bases 1 to 118788)
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,B., Allen,N.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Melidrm,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
Zody,M.

TITLE
JOURNAL
Direct Submission
Submitted (10-JUL-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS
3 (bases 1 to 118788)
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,B., Allen,N.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatn,C.,
Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Hagos,B.,
Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Melidrm,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Nachman,A., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B.,
Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
Stange-Thomann,N., Stiwell,J., Stojanovic,N., Stone,C.,
Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H.,
Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and
Zody,M.

TITLE
JOURNAL
Direct Submission
Submitted (23-JUL-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Jul 23, 1998 this sequence version replaced gi:3335015.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES
source
1. 118788
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="hrpk:597 M 12"
/clone.lib="RPCI-11 human BAC library"
795..884
/rpt_family="MIR"
complement(932..1000)
/rpt_family="MIR"
2044..2076
/rpt_family="AT_rich"
2237..2713
/rpt_family="MER34"
2718..3477
/rpt_family="L2"
complement(3478..3821)
/rpt_family="MLT1C"
3822..4007
/rpt_family="L2"
complement(4015..4281)
/rpt_family="MER89"
complement(4282..4321)
/rpt_family="(TAAA)n"
complement(4328..4608)
/rpt_family="AluSx"
complement(4609..4895)
/rpt_family="MER89"
complement(4896..5186)
/rpt_family="AluSp"
5219..5479
/rpt_family="L2"
complement(5481..5740)
/rpt_family="MER34"
complement(5963..6275)
/rpt_family="AluSp"
complement(6411..6494)
/rpt_family="MER5A"
6830..7013
/rpt_family="MLT1C"
7014..7318
/rpt_family="AluYb8"
7319..7581
/rpt_family="MLT1C"
7721..8032
/rpt_family="AluJb"
8039..8333
/rpt_family="AluSx"
9431..9460
/rpt_family="AT_rich"
9662..9829
/rpt_family="MIR"
complement(11302..11728)
/rpt_family="MER39"
complement(11824..11884)
/rpt_family="LTR16C"
complement(11928..12108)
/rpt_family="LTR16C"
12453..12725
/rpt_family="AluJb"
13377..13785
/rpt_family="MSTB"
14118..14153
/rpt_family="(CA)n"
complement(15320..15620)
/rpt_family="AluJo"
15683..15759
/rpt_family="(CA)n"
complement(15838..15893)
/rpt_family="MER91A"
complement(15915..16080)
/rpt_family="MER91A"
complement(16474..16544)
/rpt_family="L2"
17907..17920
/rpt_family="L1ME"
18271..18311
/rpt_family="L2"
complement(18368..18529)
/rpt_family="AluJb"
complement(18530..18560)
/rpt_family="(CA)n"
complement(18561..18681)
/rpt_family="AluJb"
18760..19065
/rpt_family="AluSg"
complement(19638..19896)
/rpt_family="MIR"
complement(19971..20260)
/rpt_family="L1ME3A"
20263..20552
/rpt_family="MER58B"
complement(20553..20816)
/rpt_family="L1ME3A"
complement(20821..21069)
/rpt_family="L1M4"
complement(21071..21100)
/rpt_family="(TAAA)n"

```

```
repeat_region complement(21101..21338)
repeat_region /rpt_family="AluSx"
repeat_region complement(21348..21598)
repeat_region /rpt_family="L1M4"
repeat_region 21799..22104
repeat_region /rpt_family="AluJo"
repeat_region 22151..22562
repeat_region /rpt_family="MSTA"
repeat_region 23978..24287
repeat_region /rpt_family="AluSq"
repeat_region 24852..25199
repeat_region /rpt_family="L2"
repeat_region complement(25697..25985)
repeat_region /rpt_family="MUT1D"
repeat_region 25816..25895
repeat_region /note="Single-stranded coverage."
repeat_region complement(25994..26099)
repeat_region /rpt_family="purine-rich"
repeat_region complement(26122..26291)
repeat_region /rpt_family="MUT1D"
repeat_region 26312..26337
repeat_region /rpt_family="(CAAAA)n"
repeat_region 26456..26480
repeat_region /rpt_family="(CGGAA)n"
repeat_region complement(26596..26781)
repeat_region /rpt_family="MER6B"
repeat_region complement(26912..26990)
repeat_region /rpt_family="MER21B"
repeat_region 26991..27290
repeat_region /rpt_family="AluSx"
repeat_region complement(27231..27975)
repeat_region /rpt_family="MER21B"
repeat_region 28034..28197
repeat_region /rpt_family="LTR28"
repeat_region 28266..28795
repeat_region /rpt_family="LTR28"
repeat_region 28843..28916
repeat_region /rpt_family="LTR28"
repeat_region 28920..29005
repeat_region /rpt_family="MIR"
repeat_region 29140..29233

Query Match 100.0%; Score 500; DB 9; Length 118788;
Best Local Similarity 100.0%; Pred. NO. 1.le-116;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATCAGGTGACTGAATTCATATATCTGAAGTAGGAGATCTGTTATTGCTGTTATTACAT 60
Db 77653 TATCAGGTGACTGAATTCATATATCTGAAGTAGGAGATCTGTTATTGCTGTTATTACAT 77712

QY 61 TTTACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGAGCTAACAAATGAGCC 120
Db 77713 TTTACATAAGAAAGCTGAGGCTCTGAGAGGTCAAGATCACGAGCTAACAAATGAGCC 77772

QY 121 AGACTCTTGCTTTAGAGCTTGCTCTATCTGCTTTTCTTTTCCAAAACACTACAA 180
Db 77773 AGACTCTTGCTTTAGAGCTTGCTCTATCTGCTTTTCTTTTCCAAAACACTACAA 77832

QY 181 TTTTCTGTTTGTGTTTGTGTTTTCAGACAGGCTCGAGGTTCCACCAGGCTGAGCT 240
Db 77833 TTTTCTGTTTGTGTTTGTGTTTTCAGACAGGCTCGAGGTTCCACCAGGCTGAGCT 77892

QY 241 GCAGTGGCGGATTTGCACTCACCGCAACCTCCGCTCCGCGTTTAAAGCATTCCTCTGC 300
Db 77893 GCAGTGGCGGATTTGCACTCACCGCAACCTCCGCTCCGCGTTTAAAGCATTCCTCTGC 77952

QY 301 CTGAGCTCCCAAGTAGTGGGACTCAAGCTCGGACACACCGTAAAAATGATCAAGTT 360
Db 77953 CTGAGCTCCCAAGTAGTGGGACTCAAGCTCGGACACACCGTAAAAATGATCAAGTT 78012

QY 361 CTAAACATGATGATACGATACGATTAACATGGAATAAAATTAGCAACGCGTTATGCTAATG 420
Db 78013 CTAAACATGATGATACGATTAACATGGAATAAAATTAGCAACGCGTTATGCTAATG 78072
```

```
QY 421 CTCAATACAATTGATTCTCTACATTTAATCTCTCAACCACTACAACCACTCTTAATCTC 480
Db 78073 CTCAATACAATTGATTCTCTACATTTAATCTCTCAACCACTACAACCACTCTTAATCTC 78132

QY 481 AGCTCTCTGAGGACTGACGT 500
Db 78133 AGCTCTCTGAGGACTGACGT 78152

RESULT 2
AC026279 157159 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 15 clone RP11-321G12, WORKING DRAFT
DEFINITION AC026279
ACCESSION AC026279
VERSION AC026279.4 GI:8954310
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 157159)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 157159)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Genome Sequencing Center, Washington
MO 63108, USA
COMMENT On Jul 7, 2000 this sequence version replaced gi:7523953.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0312G12
----- Summary Statistics -----
Sequencing vector: M13; 86%
Sequencing method: plasmid; 14%
Chemistry: Dye-primer ET; 86% of reads
Chemistry: Dye-terminator Big Dye; 14% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148266 bases at least Q40
Consensus quality: 151490 bases at least Q30
Consensus quality: 153444 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 156059; sum-of-contigs
Quality coverage: 4.15 in Q20 bases; agarose-fp
Quality coverage: 3.96 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1812: contig of 1812 bp in length
* 1813: gap of unknown length
* 1913: contig of 2138 bp in length
* 4050: gap of unknown length
* 4151: contig of 4742 bp in length
* 8892: gap of unknown length
* 8893: contig of 5346 bp in length
* 14338: gap of unknown length
* 14439: contig of 5622 bp in length
* 20061: gap of unknown length
* 20161: contig of 7084 bp in length
```

* 27245 27344: gap of unknown length
* 27345 42412: contig of 15068 bp in length
* 42413 42512: gap of unknown length
* 42513 55808: contig of 13296 bp in length
* 55809 55908: gap of unknown length
* 55909 69598: contig of 13690 bp in length
* 69599 69698: gap of unknown length
* 69699 88164: contig of 18466 bp in length
* 88165 88264: gap of unknown length
* 88265 106938: contig of 18674 bp in length
* 106939 107039: gap of unknown length
* 107039 157159: contig of 50121 bp in length.

FEATURES

Location/Qualifiers

Source

1..157159

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

/clone="RP11-321G12"

1..1812

/note="assembly_name:Contig1"

1913..4050

/note="assembly_name:Contig2"

4151..8892

/note="assembly_name:Contig3"

8993..14338

/note="assembly_name:Contig4"

14439..20060

/note="assembly_name:Contig5"

20161..27244

/note="assembly_name:Contig6"

27345..42412

/note="assembly_name:Contig7"

42513..55808

/note="assembly_name:Contig8"

55909..69598

/note="assembly_name:Contig9"

69699..88164

/note="assembly_name:Contig10"

88265..106938

/note="assembly_name:Contig11"

clone_end:SP6

vector_side:left

107039..157159

/note="assembly_name:Contig12"

clone_end:T7

vector_side:left

BASE COUNT 42295 a 34576 c 35453 g 43728 t 1107 others

ORIGIN

Query Match 26.2%; Score 131; DB 2; Length 157159;
Best Local Similarity 63.3%; Pred. No. 4.2e-23;
Matches 217; Conservative 0; Mismatches 125; Indels 1; Gaps 1;
QY 13 GAATTCATATCTCGAAGTAGGATGATCTGTTATTCGTTGTTATTCATTTACATTAAG 72
|||||
Db 147633 GAAAAATATTTGAGACTCCCAAGACGACGAGATATTATACAGATATATAAGAA 147692
QY 73 AAAGCTGAGGCTCTGAGAGGTCAAGATCAGCAGGTAAACAATAGCCCAAGACTTTGCT 132
Db 147693 AGTCCTCAAAATCAATCAGAAAAAGAGAATAACAAGGAAAAATGGCAAGGATATGAATA 147752
QY 133 TTAGAGCTGTGCTCTATCTGTTCTTTCTTCCAAAAAACAACACTATTTTGTGTTGT 192
Db 147753 TGTGACTCACCACCAAAAAACCCACTGGTTTGTCTCAACCTCACTAGTTTGTGTTTGT 147812
QY 193 TTTGTTTGTGTTGAGACAGGCTCTCGAGGTGTCCAGGCTGAGTGCAGTGCAGCGCGCA 252
Db 147813 TTTGTTTGTGTTGAGACAGAGCTCTCACTGTGTCACGAGCTGAGTGCAGTGCAGCA 147872
QY 253 TTTCGACTCACCGCAACCTCCGCTCC -GCGCTTAAAGGATTTCTCTGCTCAGCTCCC 311
Db 147873 TCTCGGCTCACTGCACACCTCCGCTCTCTGGGTTAAAGTGAATTCCTGCTCAGCTCCC 147932

QY 312 AAGTAGCTGGGACTACAAAGCTCGGGACACACACGTAATAATGAT 354
|||||
Db 147933 GAGTAGCTGGGACTACAGATGCCCGCCACACGCCCACTAAT 147975

RESULT 3

AC087525/c

LOCUS

AC087525 157590 bp DNA linear PRI 14-MAY-2002
Homo sapiens chromosome 15, clone RP11-321G12, complete sequence.

DEFINITION

AC087525

ACCESSION

AC087525.6

VERSION

GI:20564446

KEYWORDS

HTG.

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 15, clone RP11-321G12

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,

Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,

Collymore,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S.,

Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,

Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,

Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,

Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,

Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,

Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,

Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Travers,M., Travis,N., Triglio,J., Vassiliev,H., Viel,R., Vo,A.,

Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,

Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

3 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

4 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

4 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

4 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

4 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

REFERENCE

4 (bases 1 to 157590)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAvellino,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagoes, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (14-MAY-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 14, 2002 this sequence version replaced gi:20334597.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center -----
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L11919
 Center clone name: 321_G12

TITLE JOURNAL COMMENT

FEATURES

source

Location/Qualifiers

1..157590
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-321G12"
 /clone_lib="RPC1-11 Human Male BAC"
 complement(177..301)
 /rpt_family="L3"
 302..598
 /rpt_family="AluJb"
 495..499
 /note="<30 qual SNGL region"
 550..560
 /note="<30 qual SNGL region"
 complement(599..940)
 /rpt_family="L3"
 676..681
 /note="<30 qual SNGL region"
 721..993
 /note="single clone coverage"
 969..985
 /rpt_family="(A)n"
 994..1000
 /note="<30 qual SNGL region"
 1156..1179
 /note="<30 qual SNGL region"
 1195..1228
 /rpt_family="AT_rich"
 1238..1245
 /note="<30 qual SNGL region"
 1275..1283
 /note="<30 qual SNGL region"
 1343..1349
 /note="<30 qual SNGL region"

2197..2204
 /note="<30 qual SNGL region"
 2216..2220
 /note="<30 qual SNGL region"
 2239..2243
 /note="<30 qual SNGL region"
 2325..2335
 /note="<30 qual SNGL region"
 2590..2625
 /rpt_family="AT_rich"
 complement(2683..2961)
 /rpt_family="AluJb"
 complement(2969..3579)
 /rpt_family="L1M4"
 complement(3580..3652)
 /rpt_family="Tigger3(Golem)"
 complement(3653..3947)
 /rpt_family="AluJb"
 complement(3948..4149)
 /rpt_family="Tigger3(Golem)"
 complement(4152..4212)
 /rpt_family="L1PA13"
 complement(4211..4279)
 /rpt_family="L1PA15"
 4848..4886
 /rpt_family="AT_rich"
 complement(5070..5189)
 /rpt_family="FLAM_C"
 8698..8741
 /rpt_family="AT_rich"
 8918..9153
 /rpt_family="MER46B"
 9154..9455
 /rpt_family="AluJb"
 9456..9578
 /rpt_family="MER46B"
 10133..10286
 /rpt_family="L2"
 complement(10284..10408)
 /rpt_family="L2"
 10484..10588
 /rpt_family="MIR"
 10905..11082
 /rpt_family="L3"
 complement(11814..12123)
 /rpt_family="AluJb"
 12577..12610
 /rpt_family="(TTTTA)n"
 complement(12615..12899)
 /rpt_family="AluY"
 complement(12920..13194)
 /rpt_family="AluSg"
 complement(13202..13466)
 /rpt_family="AluJb"
 14372..14498
 /rpt_family="L2"
 14833..14852
 /rpt_family="(TTTTG)n"
 14891..14984
 /rpt_family="L3"
 complement(14985..15197)

Query Match 26.2%; Score 131; DB 9; Length 157590;

Best Local Similarity 63.3%; Pred. No. 4.2e-23;

Matches 217; Conservative 0; Mismatches 125; Indels 1; Gaps 1;

Qy 13 GAATTCATATTTCTGAAGTAGGAGATCTGTTATTGCTGTTATTACATTTTACACATAAG 72

Db 116995 GAAAAAATAATTGAGACTCCCAAGACCAGCAAGAGATTATTATACAGATATATAAGAA 116936

Qy 73 AAAGCTGAGGCTCTGAGAGGTCAAGATCAGCAGCTACAATGAGCCCAAGACTCTTGGT 132

Db 116935 AGTCCTACAAATCAATCAGAAAAAGAGAATACACAGAAAAATGGCAAGGATATGAATA 116876

RESULT 5
AC017038
LOCUS
DEFINITION
20 unordered pieces.
AC017038
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 180707)
Waterston.R.H.
The sequence of Homo sapiens clone
Unpublished
REFERENCE
2 (bases 1 to 180707)
Waterston.R.H.
Direct Submission
Submitted (09-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 16, 2000 this sequence version replaced gi:7230978.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0313B02
----- Summary Statistics -----
Sequencing vector: M13; %
Chemistry: Dye-primer ET; % of reads
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168786 bases at least Q40
Consensus quality: 172458 bases at least Q30
Consensus quality: 174496 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 178807; sum-of-contigs
Quality coverage: 3.48 in Q20 bases; agarose-fp
Quality coverage: 3.86 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2076: contig of 2076 bp in length
* 2077 2176: gap of unknown length
* 2177 4769: contig of 2593 bp in length
* 4770 4869: gap of unknown length
* 4870 7558: contig of 2689 bp in length
* 7559 7658: gap of unknown length
* 7659 10945: contig of 3287 bp in length
* 10946 11045: gap of unknown length
* 11046 13945: contig of 2900 bp in length
* 13946 14046: gap of unknown length
* 14046 17245: contig of 3679 bp in length
* 17245 17825: gap of unknown length
* 17825 22157: contig of 4332 bp in length
* 22157 22257: gap of unknown length
* 22257 26143: contig of 3786 bp in length
* 26143 32663: contig of 6520 bp in length
* 32663 32763: gap of unknown length
* 32763 41257: contig of 8495 bp in length
* 41258 41358: gap of unknown length
* 41358 47568: contig of 6211 bp in length

* 47569 47668: gap of unknown length
* 47669 54391: contig of 6723 bp in length
* 54391 54491: gap of unknown length
* 54492 62210: contig of 7719 bp in length
* 62211 70009: contig of 7699 bp in length
* 70010 70109: gap of unknown length
* 70110 78968: contig of 8859 bp in length
* 78969 90881: contig of 11813 bp in length
* 90882 90981: gap of unknown length
* 90982 105415: contig of 14434 bp in length
* 105416 105516: gap of unknown length
* 105516 121842: contig of 16327 bp in length
* 121843 121943: gap of unknown length
* 121943 141937: contig of 19994 bp in length
* 141937 142037: gap of unknown length
* 142037 180707: contig of 38671 bp in length.

FEATURES

source

1..180707
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-313B2"

misc_feature
1..2076
/note="assembly_name:Contig8"
misc_feature
2177..4769
/note="assembly_name:Contig9"
misc_feature
4870..7558
/note="assembly_name:Contig10"
misc_feature
7659..10945
/note="assembly_name:Contig11"
misc_feature
11046..13945
/note="assembly_name:Contig12"
misc_feature
14046..17245
/note="assembly_name:Contig13"
misc_feature
17825..22156
/note="assembly_name:Contig14"
misc_feature
22257..26042
/note="assembly_name:Contig15"
misc_feature
26143..32662
/note="assembly_name:Contig16"
misc_feature
32763..41257
/note="assembly_name:Contig17"
misc_feature
41358..47568
/note="assembly_name:Contig18"
misc_feature
47669..54391
/note="assembly_name:Contig19"
misc_feature
54492..62210
/note="assembly_name:Contig20"
misc_feature
62311..70009
/note="assembly_name:Contig21"
misc_feature
70110..78968
/note="assembly_name:Contig22"
misc_feature
79069..90881
/note="assembly_name:Contig23"
misc_feature
90982..105415
/note="assembly_name:Contig24"
clone_end:SP6
vector_side:right
105516..121842
/note="assembly_name:Contig25"
misc_feature
121943..141936
/note="assembly_name:Contig26"
misc_feature
142037..180707
/note="assembly_name:Contig27"

BASE COUNT 51636 a 38060 c 37973 g 51122 t 1916 others
ORIGIN

Query Match 25.7%; Score 128.6; DB 2; Length 180707;
Best Local Similarity 65.5%; Pred. No. 1.7e-22;
Matches 188; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```

Qy 90 AGGTCAAGATCAGCAGCTAACAAATGAGCCAAAGACTCTTGCTTTAGAGCTTGCTCTCTA 149
Db 174248 ATGTGACTATCAGTTAAACAATAGAAAGTTATTTTCACATATTTGTAGTATGTCACAAA 174307
Qy 150 TTCTTGCTTTCTTCCAAAAAACAACACTACAAATTTTGTGTTTGTGTTTGTGTTGAGA 209
Db 174308 AACTTACAAAAACAATTTACCAAGCAACACATCATCTTTTTTTTTTTTTTTTTTTTGA 174367
Qy 210 CAGGCTCTCGAGGTGTCACCCAGGCTGGAGTCAGTGGCGGATTTGCACTCACCGCAAC 269
Db 174368 CAGAGTCTCGCTCTGTCAACCCAGGCTGGAGTGCATATGCGATCTCACTGCAAG 174427
Qy 270 CTCGCCCTCCGGCTTAAGCGATTCCTCGCTCAGCTCCCAAGTAGTGGGACTACAA 329
Db 174428 TTCCGCTTCCCGGTTCTGCCATTTCTTCGCTCAGCTCCCAAGTAGTGGGACTACAG 174487
Qy 330 GTCGGGACACCACTGAAAAATGATCAAGTTCTTAACATGATGATGCATA 376
Db 174488 GCACCTGCCACCAACCCAGCTAAATTTTTTTGTATATTTTTTAGTA 174534

RESULT 6
AC060764
LOCUS Homo sapiens chromosome 8, clone RP11-626A5, complete sequence.
ACCESSION AC060764
VERSION AC060764.13 GI:22382034
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 184338)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-626A5
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 184338)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Headford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodors,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 184338)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,

```

```

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuppback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodors,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 184338)
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuppback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodors,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:22123659.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6260
Center clone name: 626_A_5
-----
Location/Qualifiers
1. 184338
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-626A5"
/clone_lib="RPC1-11 Human Male BAC"
38. .87
/rpt_family="WIR"
/complement(167 417)
/rpt_family="L1MA9"
/complement(566. 760)
/rpt_family="L1MC5"
877. .932
/rpt_family="WIR"
/complement(937. 1312)
/rpt_family="MLT1B"
1391. .1480
/rpt_family="WIR"
1481. .1794
/rpt_family="AluSq"
1919. 1926
/notes="<30 qual SNGL region"
1947. .1951
/notes="<30 qual SNGL region"
1978. .2004
/notes="single clone coverage"

```



```

repeat_region complement(2310..2617)
repeat_region /rpt_family="AluY"
repeat_region complement(2854..3158)
unsure /rpt_family="AluJb"
unsure complement(3899..3904)
repeat_region /note="<30 qual SINGL region"
repeat_region complement(3934..4089)
unsure /rpt_family="FRAM"
unsure complement(3971..3976)
unsure /note="<30 qual SINGL region"
unsure complement(4059..4063)
unsure /note="<30 qual SINGL region"
unsure complement(4122..4128)
unsure /note="<30 qual SINGL region"
unsure complement(4125)
repeat_region /note="Probably C, possibly A"
repeat_region 4640..4933
repeat_region /rpt_family="AluSg"
repeat_region complement(5011..5494)
repeat_region /rpt_family="MLT1L"
repeat_region complement(5802..5971)
repeat_region /rpt_family="MIR"
repeat_region 6791..6935
repeat_region /rpt_family="MER3"
repeat_region 6936..7241
repeat_region /rpt_family="AluSg"
repeat_region complement(8712..8915)
repeat_region /rpt_family="MIR"
repeat_region 8952..9118
repeat_region /rpt_family="MLT1B"
repeat_region 9119..9331
repeat_region /rpt_family="AluJo"
repeat_region 9360..9560
repeat_region /rpt_family="MLT1B"
repeat_region complement(9680..9801)
repeat_region /rpt_family="MER96B"
repeat_region 9802..10090
repeat_region /rpt_family="AluSc"
repeat_region complement(10091..10117)
repeat_region /rpt_family="MER96B"
repeat_region complement(10146..10355)
repeat_region /rpt_family="AluY"
repeat_region 10436..10730
repeat_region /rpt_family="AluSx"
repeat_region complement(10745..10865)
repeat_region /rpt_family="L2"
repeat_region 11163..11304
repeat_region /rpt_family="L2"
repeat_region 11305..11342
repeat_region /rpt_family="(TTTAG)n"
repeat_region complement(11343..11622)
repeat_region /rpt_family="AluJb"
repeat_region 11623..11768
repeat_region /rpt_family="L2"
repeat_region 11859..12212
repeat_region /rpt_family="THE1B"
repeat_region complement(12223..12526)
repeat_region /rpt_family="MER110"
repeat_region 12580..12707
repeat_region /rpt_family="L2"
repeat_region complement(12855..13021)
repeat_region /rpt_family="MIR"
repeat_region 13047..13466
repeat_region /rpt_family="LIMB3"
repeat_region 13467..13778
repeat_region /rpt_family="AluSg"
repeat_region 13780..13803
repeat_region /rpt_family="(TA)n"
repeat_region 13804..13854
repeat_region /rpt_family="LIMB3"
repeat_region 13881..14073
repeat_region /rpt_family="LIMB4"
repeat_region complement(14074..14373)

```

```

/rpt_family="AluJo"
14374..14663
/rpt_family="LIMB4"
14669..15128

Query Match 25.7%; Score 128.6; DB 9; Length 184338;
Best Local Similarity 65.5%; Pred. No. 1.7e-22; Mismatches 99; Indels 0; Gaps 0;
Matches 188; Conservative 0;

Qy 90 AGGTCAAGATCAGCAGCTAACAATAGACCAAGACTCTTGCTTTAGAGCTTTGCTCTCTA 149
Db 2217 ATGTGACTATCAGTTAAACAATAAGAAAGTTATTTCACATTATTGTAGTAGTCCAAA 2276

Qy 150 TCTTGCTTTCTTTCCAAAAACACTCAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 209
Db 2277 AACTTACAAAAACATTTTACCAAGCAACACATCTCTTTTCTTTTCTTTTCTTTTCTTT 2336

Qy 210 CAGGCTCTCGAGGTCTACCCAGGCTGAGTGCAGTGGCGGATTTCCGACTCACGCCAAC 269
Db 2337 CAGAGTCTCGGTCTGTCTACCCAGGCTGGAGTGCAATGATCTCAGCTCAGTGCAGAG 2396

Qy 270 CTCGCCCTCCGCGCTTAAGCGATTCTCTGCTCAGCCTCCCAAGTAGTGGGACTACAA 329
Db 2397 TTCGCCCTCCGCGGTTCTGCCATTCTCTGCTCAGCCTCCCAAGTAGTGGGACTACAG 2456

Qy 330 GCTCGGGACACACCGTAAATGATCAAGTTCTAATCATGTATGCATA 376
Db 2457 GCACCTGCCACACACCCAGCTAATTTTTTTGTATATTTTGTATA 2503

RESULT 7
AC009771_2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AC009771 Accession AC009771
Fragment Name Begin End
AC009771_0 1 110000
AC009771_1 100001 210000
AC009771_2 200001 310000
AC009771_3 300001 353131
Continuation (3 of 4) of AC009771 from base 200001 (AC009771 Homo sapiens chromosome 12

Query Match 25.3%; Score 126.4; DB 2; Length 110000;
Best Local Similarity 77.0%; Pred. No. 6.4e-22; Mismatches 46; Indels 0; Gaps 0;
Matches 154; Conservative 0;

Qy 145 CTCATTCTTGCTTTCTTTCCAAAAACACTACAAATTTTGTGTTTGTGTTTGTGTTTGT 204
Db 65251 CTCAGGATGATGGCACTTAACAAAAGTTAGTTCTTTTCTTTTCTTTTCTTTTCTTT 65192

Qy 205 TGAGACAGGCTCTCGAGGTGTCAACCCAGGCTGGAGTGCAGTGGCGCGATTTTCGACTCACC 264
Db 65191 TTAGACAGAACCTTGCCTGTCAACCCAGGCTGGAGTGCATGGCACGATATTGACTCACT 65132

Qy 265 GCAACCTCGGCTCCGCGCTTAAGCGATTCTCTGCTCAGCTCCCAAGTAGTGGGAC 324
Db 65131 GCAACCTCGGCTCCGAGTTCAAGAGATTCTCTGCTCAGCTCCCGAGTAGTGGGAC 65072

Qy 325 TACAAGCTCGGACACACACG 344
Db 65071 TACAGGCAAGCCACACG 65052

RESULT 8
AC012386_0
WPCOMMENT
Sequence split into 4 fragments LOCUS AC012386 Accession AC012386
Fragment Name Begin End
AC012386_0 1 110000
AC012386_1 100001 210000
AC012386_2 200001 310000
AC012386_3 300001 352304
LOCUS AC012386 352304 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-642P15, WORKING DRAFT

```

```

SEQUENCE, 24 unordered pieces.
AC012386
VERSION AC012386.36 GI:22474781
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 352304)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaraturunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaxia,J., Benton,J., Biname,G., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Huijck,S., Hume,J., Joudah,S.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,
Lorazo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Mareshwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkwo,S., Oguh,M., Okwuonu,G.,
Organyue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 352304)
Worley,K.C.
Direct Submission
Submitted (27-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 352304)
Worley,K.C.
Direct Submission
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 24, 2002 this sequence version replaced gi:22450376.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMPZ
Center clone name: RP11-642P15
-----
FEATURES
source
1.352304
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"

```

----- Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 342731 bases at least Q40
Consensus quality: 346690 bases at least Q30
Consensus quality: 349268 bases at least Q20
Estimated insert size: 370200; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2303: contig of 2303 bp in length
2304: gap of unknown length
2404 5961: contig of 3558 bp in length
5962 6062: gap of unknown length
6063 9865: contig of 3804 bp in length
9866 12751: gap of unknown length
12752 12852: contig of 2786 bp in length
12853 15853: gap of unknown length
15854 19977: contig of 4124 bp in length
19978 20077: gap of unknown length
20078 25949: contig of 5872 bp in length
25950 33145: gap of unknown length
33146 33245: contig of 7096 bp in length
33246 40390: gap of unknown length
40391 40489: contig of 7144 bp in length
40490 49662: gap of unknown length
49663 59828: contig of 9173 bp in length
59829 59928: gap of unknown length
59929 70187: contig of 10066 bp in length
70188 81523: gap of unknown length
81524 89620: contig of 11237 bp in length
89621 89719: gap of unknown length
89720 105265: contig of 7996 bp in length
105266 105365: gap of unknown length
105366 116344: contig of 15546 bp in length
116345 127455: gap of unknown length
127456 143190: contig of 10978 bp in length
143191 143290: gap of unknown length
143291 179464: contig of 11012 bp in length
179465 179564: gap of unknown length
179565 230543: contig of 15635 bp in length
230544 230644: gap of unknown length
230645 253493: contig of 50979 bp in length
253494 281092: contig of 22849 bp in length
281093 316138: gap of unknown length
316139 316239: contig of 27500 bp in length
316240 352304: gap of unknown length
352305 36066: contig of 34946 bp in length
36067 36066: gap of unknown length
36067 36066: contig of 36066 bp in length.

only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP11-396D18 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACEj.6

IMPORTANT: This sequence is not the entire insert of clone RP11-396D18. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
The true left end of clone RP11-175B9 is at 72090 in this sequence.
The true right end of clone RP11-572A16 is at 2000 in this sequence.

```

FEATURES             Location/Qualifiers
     source            1..74089
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-396D18"
                     /clone_lib="RP11-11.2"
BASE COUNT           20957 a 14221 c 14800 g 24111 t
ORIGIN
Query Match          25.1%; Score 125.6; DB 9; Length 74089;
Best Local Similarity 76.0%; Pred. No. 1e-21;
Matches 155; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 141 TGTCTCTATTCTGCTTTCTTCCAAAACACTACAAATTTTGTGTTGTTT 200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8257 TGTATTGTTATCATCTACTGTATTAGAAACAACTGATTTTTTTTTTTT 8198

QY 201 GTTTTGAGACAGGCTCTGAGGTGTACCCAGCTGGAGTGCAGTGGCGGATTCGACT 260
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8197 TTTTGGAGATGGAGTCTCGCTCTGTCTCCAGCTGGAGTGCAGTGGTCTTGGCT 8138

QY 261 CACCGCACTCCGCTCGCGCTTAAGGATTCCTCGCTCAGCTCCCAAGTAGCTG 320
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8137 CACTGCAACTCCGCTCCAGTTCAAGCAATTCCTGCTCAGCTCCCGAGTAGCTG 8078

QY 321 GGACTACAAGCTCGGACACACACG 344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8077 GGATTACAGGATGTGCCACCATG 8054

```

```

RESULT 14
LOCUS      CNS01DSC 208309 bp DNA linear PRI 02-MAY-2001
DEFINITION Human chromosome 14 DNA sequence BAC C-2593121 of library CalTech-D
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION  AL121767
VERSION    AL121767.6 GI:13990321
KEYWORDS   HTG.
SOURCE     human.
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 208309)
AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,

```

```

Levy M., Eckenberg, R., Bruls, T., deBerardinis, V., Craud, C.,
Gyapay, G., Saurin, W. and Weissenbach, J.
Sequencing of the human chromosome 14
Unpublished
REFERENCE      2 (bases 1 to 208309)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL       Submitted (02-MAY-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT       - Web : www.genoscope.cns.fr
              On May 7, 2001 this sequence version replaced gi:12580631.
              ----- Genom Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

```

```

-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-1074O12 (AC=AL049828)
Downstream BAC (overlapping the SP6 end) : R-34018 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 5.83x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 :
1 - 9 : 14
10 - 19 : 218
20 - 29 : 587
30 - 39 : 2543
40 - 49 : 6299
50 - 59 : 6490
60 - 69 : 10270
70 - 79 : 28018
80 - 89 : 74691
90 - 99 : 79179

```

```

Percentage of bases with a quality value >= 40 : 98 %
-----
FEATURES             Location/Qualifiers
     source            1..208309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="C-2593121"
                     /clone_lib="CalTech-D"
                     /notes="matching EMBL:G04409
                     RHdb:RH53476
                     RHdb:RH36333
                     dbSTS:STS25812
                     Identified using the e-PCR software (G. Schuler)"
BASE COUNT           70845 a 38028 c 36323 g 63113 t
ORIGIN
Query Match          25.1%; Score 125.6; DB 9; Length 208309;
Best Local Similarity 80.0%; Pred. No. 9.6e-22;
Matches 160; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

```

```

QY 146 TCTATTCTGCTTTTCTTCCAAAACACTACAAATTTTGTGTTGTTTGT 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74302 TTTCTTATCTACATCTCTACCAAAATCATAAATTTTTTTTTTTTTTTT 74243

QY 206 GAGACAGGCTCTCGAGGTGTACCCAGCTGGAGTGCAGTGGCGGATTCGACTCACCG 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74242 GAGACGGAGTCTCCCTGTGTCTCCAGCTGGAGTGCAGTGGCGGATTCGACTCACCG 74183

QY 266 CAACCTCGCTCCCGCTT-AAGCGATTCTCTGCTCAGCTCCCAAGTAGCTGGAC 324
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74182 CAAGCTCGCTCCCGAGTTCACGCCATTCTCTGCTCAGCTCCCAAGTAGCTGGAC 74123

QY 325 TACAAGCTCGGACACACACG 344

```

```

Db 74122 TACAGGCACCTGCACCATG 74103
|||||
RESULT 15
AC073228
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-660B16, WORKING DRAFT
ACCESSION AC073228
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 153185)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 153185)
Waterston,R.H.
Direct Submission
Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 27, 2000 this sequence version replaced gi:8440040.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0660B16
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer ET; 100% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 133295 bases at least Q40
Consensus quality: 138801 bases at least Q30
Consensus quality: 141663 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 149285; sum-of-contigs
Quality coverage: 2.76 in Q20 bases; agarose-fp
Quality coverage: 3.17 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1
1119 1118: contig of 1118 bp in length
1219 1218: gap of unknown length
1219 2910: contig of 1692 bp in length
2911 3010: gap of unknown length
3011 4398: contig of 1388 bp in length
4399 4499: gap of unknown length
4499 5572: contig of 1074 bp in length
5573 5672: gap of unknown length
5673 7062: contig of 1390 bp in length
7063 7163: gap of unknown length
7163 8864: contig of 1702 bp in length
8865 8965: gap of unknown length
8965 10477: contig of 1512 bp in length
10477 12553: gap of unknown length
12553 12557: contig of 1977 bp in length
12557 12654: gap of unknown length
12654 16052: contig of 3399 bp in length

FEATURES
Source
1. 153185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-660B16"
1. 1118
/note="assembly_name:Contig19"
1219..2910
/note="assembly_name:Contig20"
3011..4398

misc_feature
1. 1118
/note="assembly_name:Contig19"
misc_feature
1219..2910
/note="assembly_name:Contig20"
misc_feature
3011..4398

```

```

* 16053 16152: gap of unknown length
* 16153 18644: contig of 2492 bp in length
* 18645 18744: gap of unknown length
* 18745 20549: contig of 1805 bp in length
* 20550 20649: gap of unknown length
* 20650 22489: contig of 1840 bp in length
* 22490 22589: gap of unknown length
* 22590 25620: contig of 3031 bp in length
* 25621 25720: gap of unknown length
* 25721 28331: contig of 2611 bp in length
* 28332 28431: gap of unknown length
* 28432 31684: contig of 3253 bp in length
* 31685 31785: gap of unknown length
* 31785 35445: contig of 3661 bp in length
* 35446 38112: contig of 2567 bp in length
* 38113 38212: gap of unknown length
* 38213 41759: contig of 3547 bp in length
* 41760 41859: gap of unknown length
* 41860 44973: contig of 3114 bp in length
* 44974 45074: gap of unknown length
* 45074 49062: contig of 3988 bp in length
* 49062 52768: contig of 3607 bp in length
* 52769 52868: gap of unknown length
* 52869 55849: contig of 2981 bp in length
* 55850 59672: contig of 3723 bp in length
* 59673 59773: gap of unknown length
* 59773 62557: contig of 2785 bp in length
* 62558 62658: gap of unknown length
* 62659 66220: contig of 3562 bp in length
* 66220 70085: gap of unknown length
* 70086 70186: gap of unknown length
* 70186 73879: contig of 3694 bp in length
* 73880 73979: gap of unknown length
* 73980 79836: contig of 5757 bp in length
* 79837 79837: gap of unknown length
* 79837 84227: contig of 4291 bp in length
* 84228 84227: gap of unknown length
* 84228 85206: contig of 4979 bp in length
* 85207 89306: gap of unknown length
* 89307 94087: contig of 4781 bp in length
* 94088 94187: gap of unknown length
* 94188 99446: contig of 5259 bp in length
* 99447 99547: gap of unknown length
* 99547 104310: contig of 4764 bp in length
* 104311 104410: gap of unknown length
* 104411 110222: contig of 5812 bp in length
* 110223 110322: gap of unknown length
* 110323 116064: contig of 5742 bp in length
* 116065 116164: gap of unknown length
* 116165 121387: contig of 5223 bp in length
* 121388 121487: gap of unknown length
* 121488 128499: contig of 7011 bp in length
* 128499 135537: gap of unknown length
* 135538 135637: contig of 6939 bp in length
* 135638 142005: gap of unknown length
* 142006 142105: gap of unknown length
* 142106 153185: contig of 11080 bp in length.

```

```
misc_feature /note="assembly_name:Contig22"
4499..5572
misc_feature /note="assembly_name:Contig23"
5673..7062
misc_feature /note="assembly_name:Contig24"
7163..8864
misc_feature /note="assembly_name:Contig25"
8965..10476
misc_feature /note="assembly_name:Contig26"
10577..12553
misc_feature /note="assembly_name:Contig27"
12654..16052
misc_feature /note="assembly_name:Contig28"
16153..18644
misc_feature /note="assembly_name:Contig29"
18745..20549
misc_feature /note="assembly_name:Contig30"
20650..22489
misc_feature /note="assembly_name:Contig31"
22590..25620
misc_feature /note="assembly_name:Contig32"
25721..28331
misc_feature /note="assembly_name:Contig33"
28432..31684
misc_feature /note="assembly_name:Contig34"
31785..35445
misc_feature /note="assembly_name:Contig35"
35546..38112
misc_feature /note="assembly_name:Contig36
clone_end:SP6
vector_side:left"
38213..41759
misc_feature /note="assembly_name:Contig37"
41860..44973
misc_feature /note="assembly_name:Contig38"
45074..49061
misc_feature /note="assembly_name:Contig39"
49162..52768
misc_feature /note="assembly_name:Contig40"
52869..55849
misc_feature /note="assembly_name:Contig41"
55950..59672
misc_feature /note="assembly_name:Contig42"
59773..62557
misc_feature /note="assembly_name:Contig43"
62658..66219
misc_feature /note="assembly_name:Contig44"
66320..70085
misc_feature /note="assembly_name:Contig45"
70186..73879
misc_feature /note="assembly_name:Contig46"
73980..79736
misc_feature /note="assembly_name:Contig47"
79837..84127
misc_feature /note="assembly_name:Contig48"
84228..89206
misc_feature /note="assembly_name:Contig49"
89307..94087
misc_feature /note="assembly_name:Contig50"
94188..99446
```

```
Query Match 25.1%; Score 125.4; DB 2; Length 153185;
Best Local Similarity 72.6%; Fred. No. 1.1e-21;
Matches 162; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
```

```
Qy 121 AAGACTCTTGTAGAGCTGTCTCTATCTTCTTTTCTTTTCCAAAAACACTACAA 180
Db 152615 ACGAAGCTCTGTTTCATCTGTTTTTATGCGGTGACTTTGTATGCACACTCACTC 152674
Qy 181 TTTTGTGTTTGTGTTTGTGTTTGTGAGCAGGGTCTCGAGGTGTCACCCAGGCTGGAGT 240
Rb 152675 TTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGAGTCTGCTGTCGCCAGGCTGGAGT 152734
```

```
Qy 241 GCAGTGGCGGATTTTCGACTCACGCAACCTCCGCCCTCCGGCTTAAGCGATTTCTCTGC 300
Db 152735 GTAGTGGCGGATCTCGGCTCACTGCAAGCTCTGCTCCGGTTTCAGCCGTTCTCTGC 152794
Qy 301 CTCAGCCTCCCAAGTAGCTGGGACTACAAGCTCGGGACACCAC 343
Db 152795 CTCAGGCTCTCTGAGTAGCTGGGACTACAGGCACCTGGCACAC 152837
```

```
Search completed: May 17, 2003, 21:40:29
Job time : 2183.47 secs
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 18:18:38 ; Search time 163 Seconds
(without alignments)
5565.337 Million cell updates/sec

Title: US-09-434-382-3

Perfect score: 2958

Sequence: 1 cgcggcgtaggtgaccggc.....ataaagattgagtttgcga 2958

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2958	100.0	2958	4	US-09-564-805-3
2	2874.4	97.2	2908	4	US-09-564-805-223
3	2819.6	95.3	2892	4	US-09-564-805-225
4	2481	83.9	2481	4	US-09-564-805-1
5	1645.6	55.6	2470	4	US-09-564-805-221
6	657.2	22.2	26664	4	US-09-564-805-28
7	655	22.1	655	4	US-09-564-805-27
8	297.4	10.1	350	4	US-09-564-805-210
9	295	10.0	295	4	US-09-564-805-4
10	237	8.0	238	4	US-09-328-111-315
11	145	4.9	145	4	US-09-564-805-26
12	139	4.7	139	4	US-09-564-805-16
13	139	4.7	139	4	US-09-564-805-20
14	121	4.1	121	4	US-09-564-805-24
15	120	4.1	120	4	US-09-564-805-10
16	119	4.0	119	4	US-09-564-805-18
17	118.4	4.0	326	4	US-09-564-805-212
18	113	3.8	113	4	US-09-564-805-14
19	110	3.7	110	4	US-09-564-805-22
20	100	3.4	100	4	US-09-564-805-23
21	97	3.3	97	4	US-09-564-805-19
22	96	3.2	96	4	US-09-564-805-15
23	86	2.9	86	4	US-09-564-805-17
24	79	2.7	79	4	US-09-564-805-25
25	73	2.5	73	4	US-09-564-805-13
26	71	2.4	71	4	US-09-564-805-6
27	69	2.3	69	4	US-09-564-805-9

28	65	2.2	65	4	US-09-564-805-7	Sequence 7, Appli
29	59	2.0	59	4	US-09-564-805-11	Sequence 11, Appli
30	59	2.0	59	4	US-09-564-805-12	Sequence 12, Appli
31	58.4	2.0	2517	4	US-09-315-794-51	Sequence 51, Appli
32	58.4	2.0	2517	4	US-09-389-341-51	Sequence 51, Appli
33	58	2.0	58	4	US-09-564-805-8	Sequence 8, Appli
34	51	1.7	51	4	US-09-564-805-5	Sequence 5, Appli
c 35	48.6	1.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
36	48.2	1.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
37	48.2	1.6	2580	3	US-09-050-863-2	Sequence 2, Appli
38	48.2	1.6	2580	4	US-09-359-081-2	Sequence 2, Appli
39	48.2	1.6	5452	2	US-09-130-114-1	Sequence 1, Appli
40	48.2	1.6	9600	4	US-08-910-647-1	Sequence 1, Appli
41	48.2	1.6	9600	4	US-09-620-925-1	Sequence 1, Appli
42	48.2	1.6	10596	1	US-07-884-811-15	Sequence 15, Appli
43	48.2	1.6	10596	1	US-07-885-971-15	Sequence 15, Appli
44	48.2	1.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
45	48.2	1.6	10596	1	US-08-194-088B-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1

US-09-564-805-3

; Sequence 3, Application US/09564805

; Patent No. 6333403

; GENERAL INFORMATION:

; APPLICANT: Tavtigan, Sean V.

; APPLICANT: Teng, David H.F.

; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.

; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility

; FILE REFERENCE: Gene and a Paralog and Orthologous Genes

; CURRENT APPLICATION NUMBER: US/09/564,805

; CURRENT FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: US 60/107,468

; PRIOR FILING DATE: 1998-11-06

; PRIOR APPLICATION NUMBER: 09/434,382

; PRIOR FILING DATE: 1999-11-05

; NUMBER OF SEQ ID NOS: 240

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2958

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (51)..(2531)

; OTHER INFORMATION: coding sequence as in SEQ ID NO:1

US-09-564-805-3

Query Match 100.0%; Score 2958; DB 4; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGGCTAGGTGACCGCGGCTTCTCAGTTTGTGGAGACGGCGCATGTGGGCG 60

Db 1 CGCGGGCTAGGTGACCGCGGCTTCTCAGTTTGTGGAGACGGCGCATGTGGGCG 60

Qy 61 TTTGTCTCGTGTGCGGTCCTCGCGGCGGACGACCATGTGCGAGGAGCGACCATATCG 120

Db 61 TTTGTCTCGTGTGCGGTCCTCGCGGCGGACGACCATGTGCGAGGAGCGACCATATCG 120

Qy 121 AGGACACCGCGCGCGAGCGCGCGCAAGAGACCCCTGCGGACCTGCGCACCGCGAG 180

Db 121 AGGACACCGCGCGCGAGCGCGCGCAAGAGACCCCTGCGGACCTGCGCACCGCGAG 180

Qy 181 AGAAGCGGGACCGTCGCGGGTGCTCCGGCGGCGGCAACACCGGTACCTGACGGTGGTGG 240

Db 181 AGAAGCGGGACCGTCGCGGGTGCTCCGGCGGCGGCAACACCGGTACCTGACGGTGGTGG 240

Qy 241 CAGCGGTAGCGGAGCTCGGCGCGGCTCTAGCTCTTCTCCGAGTTCAACCGGTATC 300
Db 241 CAGCGGTAGCGGAGCTCGGCGCGCGGCTCTAGCTCTTCTCCGAGTTCAACCGGTATC 300
Qy 301 TCTTCAACTGTGGAGAGCGGTTCAGAGACTCATGCGAGGACACAAGTTAAAGGTTGCTC 360
Db 301 TCTTCAACTGTGGAGAGCGGTTCAGAGACTCATGCGAGGACACAAGTTAAAGGTTGCTC 360
Qy 361 GCCTGGACAAATATCTCTGACAGAAATGCACTGTGTCTAAATGTTGGGGGCTTAAAGTGAA 420
Db 361 GCCTGGACAAATATCTCTGACAGAAATGCACTGTGTCTAAATGTTGGGGGCTTAAAGTGAA 420
Qy 421 TGATCTTACTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGACCTCCACAAAC 480
Db 421 TGATCTTACTTTAAAGAAACCGGGCTTCAAAGTGTGTAATTTCTGAGCTCCACAAAC 480
Qy 481 TGGAAAAATACCTCGAAGCAATCAAATAATTTCTGTGCTCAATGAAAGGAATAGAATGG 540
Db 481 TGGAAAAATACCTCGAAGCAATCAAATAATTTCTGTGCTCAATGAAAGGAATAGAATGG 540
Qy 541 CTGTGCGGCCCACTCTGCCCCAGAAATACGAGGATGAAACCATGACAGTTTACCAGATCC 600
Db 541 CTGTGCGGCCCACTCTGCCCCAGAAATACGAGGATGAAACCATGACAGTTTACCAGATCC 600
Qy 601 CCATACACAGTGAACAGAGGAGGGAAGCAACCACTGCGAGATCCAGAAAGGCGTC 660
Db 601 CCATACACAGTGAACAGAGGAGGGAAGCAACCACTGCGAGATCCAGAAAGGCGCGTC 660
Qy 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Db 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Qy 721 TTCCACATGTTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGGTCTGATCTTTCA 780
Db 721 TTCCACATGTTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGGTCTGATCTTTCA 780
Qy 781 TCTGTAAGCTTCACTTAAAGAGAGAAACTCTTGGTGCTCAAGCAAGAGAGATGGGCC 840
Db 781 TCTGTAAGCTTCACTTAAAGAGAGAAACTCTTGGTGCTCAAGCAAGAGAGATGGGCC 840
Qy 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCTGCTGCTCAAGCAAGGGAAGAGCA 900
Db 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCTGCTGCTCAAGCAAGGGAAGAGCA 900
Qy 901 TCACATGAAGAGAGAGAGATTTGGTGTGAAGAGCTGTACTCTCCAGATCTGTGTG 960
Db 901 TCACATGAAGAGAGAGAGATTTGGTGTGAAGAGCTGTACTCTCCAGATCTGTGTG 960
Qy 961 CTGCTTTTGTGGTGAATGTCCAGATGAAGCTTCATTCACCCCATCTGTGAGATG 1020
Db 961 CTGCTTTTGTGGTGAATGTCCAGATGAAGCTTCATTCACCCCATCTGTGAGATG 1020
Qy 1021 CCACCTTTCAAGGTACCAAGGAAGGAGATGCCCGCTGGCTTGGTGGTTACATGG 1080
Db 1021 CCACCTTTCAAGGTACCAAGGAAGGAGATGCCCGCTGGCTTGGTGGTTACATGG 1080
Qy 1081 CCCAGCATCTGTCTGTGACAGCAGGTACCAAGCTGAGTGGAGAGTTTGGGCTG 1140
Db 1081 CCCAGCATCTGTCTGTGACAGCAGGTACCAAGCTGAGTGGAGAGTTTGGGCTG 1140
Qy 1141 ACACCCAGACTTGGTCTGATGAGACTGTGCTCAGTTCAACCCCATCTGCGAGCCACA 1200
Db 1141 ACACCCAGACTTGGTCTGATGAGACTGTGCTCAGTTCAACCCCATCTGCGAGCCACA 1200
Qy 1201 AGATTCAAACCCAGCTCAACCTCAACCCCGGACATCTTCCCTGCTCACCAGTTTCC 1260
Db 1201 AGATTCAAACCCAGCTCAACCTCAACCCCGGACATCTTCCCTGCTCACCAGTTTCC 1260
Qy 1261 GCTGTGAAGAGAGGGCCCCACCTCAGTGTGCCCATGTTCAAGGTTGAATGCTCTCTCA 1320
Db 1261 GCTGTGAAGAGAGGGCCCCACCTCAGTGTGCCCATGTTCAAGGTTGAATGCTCTCTCA 1320

Qy 1321 AGTACCAGCTCCGTCCTCCAGGAGGAGTGGCAGAGGATGCATTATTACTTTGCAATCTG 1380
Db 1321 AGTACCAGCTCCGTCCTCCAGGAGGAGTGGCAGAGGATGCATTATTACTTTGCAATCTG 1380
Qy 1381 AGGAATTCATAGTTTGAAGCGCTGCGAGCTTCCCAACTTTCAGCAGAGCGTGCAGAGTACA 1440
Db 1381 AGGAATTCATAGTTTGAAGCGCTGCGAGCTTCCCAACTTTCAGCAGAGCGTGCAGAGTACA 1440
Qy 1441 GGAGGAGTGGCAGGAGCGGCCAGCCAGAGAGAAAGAGTCACTACCCAGAAATCA 1500
Db 1441 GGAGGAGTGGCAGGAGCGGCCAGCCAGAGAGAAAGAGTCACTACCCAGAAATCA 1500
Qy 1501 TCTTCTTTGAAACAGGCTCTGCCATCCCGATGAAGATTCGAAATGTTCAGTGCACACTTG 1560
Db 1501 TCTTCTTTGAAACAGGCTCTGCCATCCCGATGAAGATTCGAAATGTTCAGTGCACACTTG 1560
Qy 1561 TCAACATAAGCCCGCACACGCTCTCTGCTACTGACTGTGGTGGGGCACATTTGGGCAGC 1620
Db 1561 TCAACATAAGCCCGCACACGCTCTCTGCTACTGACTGTGGTGGGGCACATTTGGGCAGC 1620
Qy 1621 TGTGCCCTCATTTACGAGAGACAGGTGTCCTGGGCACCCCTGGCTGCTGTGTTG 1680
Db 1621 TGTGCCCTCATTTACGAGAGACAGGTGTCCTGGGCACCCCTGGCTGCTGTGTTG 1680
Qy 1681 TGTCCCACTGCGAGGATCACACACGGGCTTGCCAACTATCTTCTGCTGCAGAGAGAAC 1740
Db 1681 TGTCCCACTGCGAGGATCACACACGGGCTTGCCAACTATCTTCTGCTGCAGAGAGAAC 1740
Qy 1741 GCGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTTCTGGTGGTGGCCCAACCCAGC 1800
Db 1741 GCGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTTCTGGTGGTGGCCCAACCCAGC 1800
Qy 1801 TCAAGAGCTGGCTCCAGCAGTACACAAACAGTGCAGAGGCTCTGCACACATCAGTA 1860
Db 1801 TCAAGAGCTGGCTCCAGCAGTACACAAACAGTGCAGAGGCTCTGCACACATCAGTA 1860
Qy 1861 TGATTCCTGCCAAATGCTTTCAGGAGGCTGAGATCTCCAGTCTCCAGTGGAAAGAT 1920
Db 1861 TGATTCCTGCCAAATGCTTTCAGGAGGCTGAGATCTCCAGTCTCCAGTGGAAAGAT 1920
Qy 1921 TGATCAGTTCCGCTTTCGGAACATGTATTTGGAAGAGTTTCAGACCTGTCTGGTGGCGC 1980
Db 1921 TGATCAGTTCCGCTTTCGGAACATGTATTTGGAAGAGTTTCAGACCTGTCTGGTGGCGC 1980
Qy 1981 ACTGCAAGCATGCTTGGCTGTGGCTGTGCACACCTCTGGCTGGAAGTGTCTATT 2040
Db 1981 ACTGCAAGCATGCTTGGCTGTGGCTGTGCACACCTCTGGCTGGAAGTGTCTATT 2040
Qy 2041 CCGGGACACCATGCGCTCGAGGCTCTGTCGGATGGGAAAGATGCCACCCCTCTCTGA 2100
Db 2041 CCGGGACACCATGCGCTCGAGGCTCTGTCGGATGGGAAAGATGCCACCCCTCTCTGA 2100
Qy 2101 TACATGAAGCCACCTTGAAGAGTGGTTTGAAGAGGAGCAGTGGAAAGACACACAGCA 2160
Db 2101 TACATGAAGCCACCTTGAAGAGTGGTTTGAAGAGGAGCAGTGGAAAGACACACAGCA 2160
Qy 2161 CAACGTCCTCAAGCCATCAGCTGGGATCGGATGAACGCGGAGTTCATTATGCTGAACC 2220
Db 2161 CAACGTCCTCAAGCCATCAGCTGGGATCGGATGAACGCGGAGTTCATTATGCTGAACC 2220
Qy 2221 ACTTCAGCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAGAGTGG 2280
Db 2221 ACTTCAGCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGCGAGAGAGTGG 2280
Qy 2281 GAGTTGCTTTTGAACCATGAAGTCTGCTTTTGGAGACTTTCACCAATGCCCAAGCTGA 2340
Db 2281 GAGTTGCTTTTGAACCATGAAGTCTGCTTTTGGAGACTTTCACCAATGCCCAAGCTGA 2340
Qy 2341 TTCCCCCAGCTGAAGCCCTGTTTCTGGCGACATCGAGGAGATGGAGGCGGAGGAGA 2400
Db 2341 TTCCCCCAGCTGAAGCCCTGTTTCTGGCGACATCGAGGAGATGGAGGCGGAGGAGA 2400
Qy 2401 AGCGGAGCTGCGCAGGTGCGGCGGCCCTCTCTGTCCAGGAGCTGGCAGGCGCCTGG 2460

```
Db 2401 AGCGGAGCTGCGGAGGTCGGGCGGCCCTCTCTGTCAGGAGCTGGCAGGCGCTGG 2460
Qy 2461 AGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGGACACAGGCAAGAGGTCA 2520
Db 2461 AGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGGACACAGGCAAGAGGTCA 2520
Qy 2521 GAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTGCCCCAG 2580
Db 2521 GAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTTCTGCCCCAG 2580
Qy 2581 CACGCAACCGTATCTGCCCTCTCTGCTGTAGAGCTGAAGAGCGGTCCGCCAGGAGG 2640
Db 2581 CACGCAACCGTATCTGCCCTCTCTGCTGTAGAGCTGAAGAGCGGTCCGCCAGGAGG 2640
Qy 2641 CAGCTCAGATAGGTGATGAGGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Db 2641 CAGCTCAGATAGGTGATGAGGCTGTGCCGAGGCTTGGGCTCCACATAAGCACTAGT 2700
Qy 2701 CTATAGATGCTCTTAGGACTGGTGGCAGACGCGGGCCAGGAGGCTGCCACAGG 2760
Db 2701 CTATAGATGCTCTTAGGACTGGTGGCAGACGCGGGCCAGGAGGCTGCCACAGG 2760
Qy 2761 GAAGCAAGCAGATGAACATAATTTCAATTCAGGCGCTTTTAAAGAGTCTTTGGAAACAG 2820
Db 2761 GAAGCAAGCAGATGAACATAATTTCAATTCAGGCGCTTTTAAAGAGTCTTTGGAAACAG 2820
Qy 2821 ACGGGGCACTTTCTCTTAATCCAGCAAGTGATTCCTCGCACACAGCAGAGCAAGCAGA 2880
Db 2821 ACGGGGCACTTTCTCTTAATCCAGCAAGTGATTCCTCGCACACAGCAGAGCAAGCAGA 2880
Qy 2881 GTAACAGGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATTTCAGCTGCAA 2940
Db 2881 GTAACAGGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATTTCAGCTGCAA 2940
Qy 2941 TAAAGATTGAGTTTGCAA 2958
Db 2941 TAAAGATTGAGTTTGCAA 2958
```

RESULT 2

```
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-564-805-223
```

Query Match 97.2%; Score 2874.4; DB 4; Length 2908;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 2887; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy 51 ATGTGGGCGCTTTGCTCCTCTGCTGCGGTCCGGCGCCGACGCACTATCTCGAGGAGCGC 110
Db 1 ATGTGGGCGCTTTGCTCCTCTGCTGCGGTCCGGCGCCGACGCACTATCTCGAGGAGCGC 60
Qy 111 ACCATATCGCAGGACCCCGCGCGCGAGCGCGCCGACGAGGACCCGCTCGGCGACCTG 170
Db 61 ACCATATCGCAGGACCCCGCGCGCGAGCGCGCCGACGAGGACCCGCTCGGCGACCTG 120
Qy 171 CGCAGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGGCCCAAAACACCGTGTACCTG 230
Db 121 CGCAGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGGCCCAAAACACCGTGTACCTG 180
Qy 231 CAGGTGTGGCAGCGGCTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCAGTTTC 290
Db 181 CAGGTGTGGCAGCGGCTAGCCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCAGTTTC 240
Qy 291 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTCAGAGACTCATGCGAGGACCAAGTTA 350
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTCAGAGACTCATGCGAGGACCAAGTTA 300
Qy 351 AAGGTGTCTCGCCTGGGCAACATATTCCTGACAGCAATGCACTGTCTAAATGTTGGGGC 410
Db 301 AAGGTGTCTCGCCTGGGCAACATATTCCTGACAGCAATGCACTGTCTAAATGTTGGGGC 360
Qy 411 TTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Db 361 TTAAGTGAATGATTTCTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Qy 471 CCTCCACAACCTGGGAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTTGAAGA 530
Db 421 CCTCCACAACCTGGGAAATACCTCGAGCAATCAAAATATTTCTGGTCCATTTGAAGA 480
Qy 531 ATAGAACTGGGTGTGGCGGCCCACTCTGCCCGAGAAATACGAGGATGAACCAATGACAGT 590
Db 481 ATAGAACTGGGTGTGGCGGCCCACTCTGCCCGAGAAATACGAGGATGAACCAATGACAGT 540
Qy 591 TACCAGATCCCCTATACAGTGAACAGAGGAGGAAAGCAACCAATGCGAGTCCA 650
Db 541 TACCAGATCCCCTATACAGTGAACAGAGGAGGAAAGCAACCAATGCGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATCTTCAGACTCCGAGTCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGGATCTTCAGACTCCGAGTCGAATGAAAT 660
Qy 711 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGAGGCTCAGGACTCTTCCCTGGTC 770
Db 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGAGGCTCAGGACTCTTCCCTGGTC 720
Qy 771 GTAGCTTTTCTGCTGCTGCTTCACTTTAAAGAGAGGAAACTTCTTGGTGTCTCAAAAG 830
Db 721 GTAGCTTTTCTGCTGCTGCTTCACTTTAAAGAGAGGAAACTTCTTGGTGTCTCAAAAG 780
Qy 831 GAGATGGGCTCCAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 890
Db 781 GAGATGGGCTCCAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 891 GGGAAAGCATTCTCATGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 950
Db 841 GGGAAAGCATTCTCATGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Qy 951 GATCCTGTGTCTGCTTTTGTGGTGGTGAATGTCCAGATGAAGGCTTCATTTCAACCCATC 1010
Db 901 GATCCTGTGTCTGCTTTTGTGGTGGTGAATGTCCAGATGAAGGCTTCATTTCAACCCATC 960
Qy 1011 TGTGAGATGCCCTTTTCAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1070
Db 961 TGTGAGATGCCCTTTTCAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Qy 1071 GTTCACATGGGCGCCAGCATCTGCTGTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1130
Db 1021 GTTCACATGGGCGCCAGCATCTGCTGTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
```

QY 1131 TTGGGCTGACCCAGCACTTGGTCTGAATGAGACTGTGCTCAGTTTCAACCTT 1190
Db 1081 TTGGGCTGACCCAGCACTTGGTCTGAATGAGAACTGTGCCCTCAGTTTCAACCTT 1140
QY 1191 CGCAGCCACAAGATTTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCCCTGCTC 1250
Db 1141 CGCAGCCACAAGATTTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCCCTGCTC 1200
QY 1251 ACCAGTTTCCGCTGTAAGAGAGAGGCCCCACCTCTCAGTGTGCCCATGGTTCAGGGTAA 1310
Db 1201 ACCAGTTTCCGCTGTAAGAGAGAGGCCCCACCTCTCAGTGTGCCCATGGTTCAGGGTAA 1260
QY 1311 TGCCTCTCAAGTACCAGCTCCGTCAGAGAGGAGTGGCAGAGGATGCCATTATTACT 1370
Db 1261 TGCCTCTCAAGTACCAGCTCCGTCAGAGAGGAGTGGCAGAGGATGCCATTATTACT 1320
QY 1371 TGCATCTCAGGAAATTCATAGTTGAGCGCTGCACTTCCCACTTCCAGCAGAGCGTG 1430
Db 1321 TGCATCTCAGGAAATTCATAGTTGAGCGCTGCACTTCCCACTTCCAGCAGAGTG 1380
QY 1431 CAGAGTACAGAGAGGTGCGCAGGACCGCCAGCCAGCCAGCAGAGAGAGAAAGTCACTAC 1490
Db 1381 CAGAGTACAGAGAGGTGCGCAGGACCGCCAGCCAGCCAGCAGAGAGAGAAAGTCACTAC 1440
QY 1491 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATCCAGATTCGAATGTCAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATTCGAATGTCAGT 1500
QY 1551 GCCACATCTGTCAACATTAAGCCCCAGACAGCTCTCTGCTACTTGGACTGTGGTGAAGGACACA 1610
Db 1501 GCCACATCTGTCAACATTAAGCCCCAGACAGCTCTCTGCTACTTGGACTGTGGTGAAGGACAG 1560
QY 1611 TTGGGAGCTGTGCCGTCAATTAACAGAACCCAGGTGGAAGGCTCTGGGACCCCTGGCT 1670
Db 1561 TTGGGAGCTGTGCCGTCAATTAACAGAACCCAGGTGGAAGGCTCTGGGACCCCTGGCT 1620
QY 1671 GCTGTGTTTGTGTCACCTGCAACAGATCACACAGGCTTGCACAGTATCTTGGTG 1730
Db 1621 GCTGTGTTTGTGTCACCTGCAACAGATCACACAGGCTTGCACAGTATCTTGGTG 1680
QY 1731 CAGAGAGAACCGGCTTGGCATCTTTGGGAAAGCCGCTTACCCCTTCTGCTGGTGTGCC 1790
Db 1681 CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCCCTTACCCCTTCTGCTGGTGTGCC 1740
QY 1791 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGAGTCTGTCAC 1850
Db 1741 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGAGAGTCTGTCAC 1800
QY 1851 CACATCAGTATGATTCCTGCCAAATGCCCTCAGGAAGGGCTGAGATCTCCAGTCCCTGCA 1910
Db 1801 CACATCAGTATGATTCCTGCCAAATGCCCTCAGGAAGGGCTGAGATCTCCAGTCCCTGCA 1860
QY 1911 GTGGAAGATGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTCAGACCTGT 1970
Db 1861 GTGGAAGATGATCAGTTCGCTGTGGCAACATGTGATTTGGAAGAGTTCAGACCTGT 1920
QY 1971 CTGGTGGGCACTCAAGACTGCGTTTGGCTGTGCGCTGGTGCAACCTCTGGCTGGAAA 2030
Db 1921 CTGGTGGGCACTCAAGACTGCGTTTGGCTGTGCGCTGGTGCAACCTCTGGCTGGAAA 1980
QY 2031 GTGGTCTATTCGGGGACACATGCCCTCGGAGGCTCTGGTCCGATGGGAAAGATGCC 2090
Db 1981 GTGGTCTATTCGGGGACACATGCCCTCGGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
QY 2091 ACCCTCTCATATGAAGCCACCTTGAAGATGTTTGAAGAGAGAGAGTGGAAAAG 2150
Db 2041 ACCCTCTCATATGAAGCCACCTTGAAGAGAGAGTGGAAAAG 2100
QY 2151 ACACACAGACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2210
Db 2101 ACACACAGACAACTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2160
QY 2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCCACTTCAGC 2270

Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCCACTTCAC 2220
QY 2271 GAGAAAGTGGAGTGTGCTTTTGACCACATGAAGTGTGCTTTGGAGACTTTTCAACAATG 2330
Db 2221 GAGAAAGTGGAGTGTGCTTTTGACCACATGAAGTGTGCTTTGGAGACTTTTCAACAATG 2280
QY 2331 CCCAAGCTGATTTCCCCACTGAAGCCCTGTTTGTGGCCACATCGAGGAGATGGAGG 2390
Db 2281 CCCAAGCTGATTTCCCCACTGAAGCCCTGTTTGTGGCCACATCGAGGAGATGGAGG 2340
QY 2391 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCGCTCTCTGTCCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCGCTCTCTGTCCAGGAGCTGGCA 2400
QY 2451 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
QY 2511 AAGAGTCAAGGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAAGGCTGTGTCTT 2570
Db 2461 AAGAGTCAAGGCCAGTGAAGATCTGGGAGACCTTGAACCTCAGAAGGCTGTGTCTT 2520
QY 2571 CTGCCCCACACGCGCTATCTGCCCTCTTGTGGTAGAAGCTGAAGAGCAGCGTC 2630
Db 2521 CTGCCCCACACGCGCTATCTGCCCTCTTGTGGTAGAAGCTGAAGAGCAGCGTC 2580
QY 2631 CCCAGGAGCAGCTCAGGATAGTGTGAGTGTGCGGAGCTGTGCGAGGCTTGGGCTCCACAT 2690
Db 2581 CCCAGGAGCAGCTCAGGATAGTGTGAGTGTGCGGAGCTGTGCGGCTTGGGCTCCACAT 2640
QY 2691 AAGCACTAGTCTATAGATGCTCTTAGGACTGTGCTGCGCACAGCCCGGCGCAGGAG 2750
Db 2641 AAGCACTAGTCTATAGATGCTCTTAGGACTGTGCTGCGCACAGCCCGGCGCAGGAG 2700
QY 2751 CTGCCACACGGAAGCAGCAGATGAACCTAATTTCAATTTCAAGGCAGTTCCTTAAAGAAAGTC 2810
Db 2701 CTGCCACACGGAAGCAGCAGATGAACCTAATTTCAATTTCAAGGCAGTTCCTTAAAGAAAGTC 2760
QY 2811 TTGAAAACAGACCGCGCACCTTTCTCTTAATCCAGCAAAAGTGAATTCCTGTCACACAGA 2870
Db 2761 TTGAAAACAGACCGCGCACCTTTCTCTTAATCCAGCAAAAGTGAATTCCTGTCACACAGA 2820
QY 2871 GACNAGCAGAGTAACAGGATCAGTGGTCTTAAGTGTCCGAGACTTAACGAAATAGTATT 2930
Db 2821 GACNAGCAGAGTAACAGGATCAGTGGTCTTAAGTGTCCGAGACTTAACGAAATAGTATT 2880
QY 2931 TCAGCTCAATAAAGATTGAGTTTGCAA 2958
Db 2881 TCAGCTCAATAAAGATTGAGTTTGCAA 2908

RESULT 3

US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0


```
Db 1981 GTGGTCTATTCCGGGACACACATGCCCTCGAGGCTCTGGTCCGATCGGGAAGATGCC 2040
Qy 2091 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAGAGGAAGCAGTGGAAAG 2150
Db 2041 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAGAGGAAGCAGTGGAAAG 2100
Qy 2151 ACACACAGACAAAGTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2210
Db 2101 ACACACAGACAAAGTCCCAAGCCATCAGCTGGGGATGCGGATGAACCGGAGTTCAAT 2160
Qy 2211 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2270
Db 2161 ATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2220
Qy 2271 GAGAAAGTGGGAGTTCCTTTGACACATGAAGTCTCTTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTCCTTTGACACATGAAGTCTCTTTGGAGACTTTTCCAAACAATG 2280
Qy 2331 CCCAGCTGATTCCTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGGATGGAGAG 2390
Db 2281 CCCAGCTGATTCCTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGGATGGAGAG 2340
Qy 2391 CGCAGGAGAAAGCGGAGTGGCGAGGTGGCGGCGCCCTCTCTGTCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAAAGCGGAGTGGCGAGGTGGCGGCGCCCTCTCTGTCAGGAGCTGGCA 2400
Qy 2451 GCGCGCTCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGCCCAACAGAGAGCCACAGGCC 2510
Db 2401 GCGCGCTCTGGAGATGGGAGCTTCAGCAGAAAGCGGCGCCCAACAGAGAGCCACAGGCC 2460
Qy 2511 AAGAAGTTCAGAGCCAGTGAAGATCTGGGACCTGAACTCAGAAAGCTGTGTGCTT 2570
Db 2461 AAGAAGTTCAGAGCCAGTGAAGATCTGGGACCTGAACTCAGAAAGCTGTGTGCTT 2520
Qy 2571 CTGCCCCACGACGACCCGCTATCTGCCCTCTCTGCTGTGTGAAGCTGAAGACAGCTG 2630
Db 2521 CTGCCCCACGACGACCCGCTATCTGCCCTCTCTGCTGTGTGAAGCTGAAGACAGCTG 2580
Qy 2631 CCCAGGAGGAGCTCAGGATAGTGGTATGAGAGCTGTGCGAGGCTTTGGGCTCCACAT 2690
Db 2581 CCCAGGAGGAGCTCAGGATAGTGGTATGAGAGCTGTGCGAGGCTTTAGGCTCCACAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCCCTCTTAGACTGTGTGCTGCGACAGCCGCGGCGCAGG 2750
Db 2641 AAGCACTAGTCTATA-----GGTGCCTGGCACAGCCGCGGACAGGAGG 2684
Qy 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCAAGTTTTTAAAGAGTC 2810
Db 2685 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCAAGTTTTTAAAGAGTC 2744
Qy 2811 TTGGAACACAGCGCGGCACTTCTTCTTAATCCAGCAAAAGTATTCCTCGACACCCAGA 2870
Db 2745 TTGGAACACAGCGCGGCACTTCTTCTTAATCCAGCAAAAGTATTCCTCGACACCCAGA 2804
Qy 2871 GACACGAGATTAACAGATCAGTGGGTCTAAGTCTCGAGACTTAACGAAATAGTATT 2930
Db 2805 GACACGAGATTAACAGATCAGTGGGTCTAAGTCTCGAGACTTAACGAAATAGTATT 2864
Qy 2931 TCAGCTGCAATAAAGATTGAGTTTGCAC 2958
Db 2865 TCAGCTGCAATAAAGATTGAGTTTGCAC 2892
```

RESULT 4

```
US-09-564-805-1
; Sequence 1, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtighan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
```

```
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
; US-09-564-805-1
```

Query Match 83.9%; Score 2481; DB 4; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 51 ATGTGGCGCTTTGCTGCTGCTGCTGCGGTCGCGGCGGAGCGCACCATGTGCGAGGAGCG 110
Db 1 ATGTGGCGCTTTGCTGCTGCTGCTGCGGTCGCGGCGGAGCGCACCATGTGCGAGGAGCG 60
Qy 111 ACCATATCGCAGGACCCCGCGCGGAGCGCGGCGGCAAGGACCCGCTGCGGACCTG 170
Db 61 ACCATATCGCAGGACCCCGCGCGGAGCGCGGCGGCAAGGACCCGCTGCGGACCTG 120
Qy 171 CGCACGCGAGAGAGCGCGACCGTGGGGTGTCTCGGCGCGCCCAACACCGTGTACCTG 230
Db 121 CGCACGCGAGAGAGCGCGACCGTGGGGTGTCTCGGCGCGCCCAACACCGTGTACCTG 180
Qy 231 CAGGTGTGCGAGGAGTACCGGGACTCGGGGCGCGCGCTCTACGTCCTTCGAGTTC 290
Db 181 CAGGTGTGCGAGGAGTACCGGGACTCGGGGCGCGCGCTCTACGTCCTTCGAGTTC 240
Qy 291 AACCGGTATCTCTTCAACTGTGGAGAGGGGTTTCAGAGACTCATGACAGGACCAAGTTA 350
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGGGGTTTCAGAGACTCATGACAGGACCAAGTTA 300
Qy 351 AAGTTTCTCGCTGGGCAACATATTCCTGCACAGAAATGCATGGTCTAATGTTGGGGGC 410
Db 301 AAGTTTCTCGCTGGGCAACATATTCCTGCACAGAAATGCATGGTCTAATGTTGGGGGC 360
Qy 411 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Db 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Qy 471 CCTCCACAACTGGAAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 530
Db 421 CCTCCACAACTGGAAAAAATACCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 531 ATAGAACTGCTGTGGGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGCTGTGGGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAGGACCAACCATCGCAGAGTCCA 650
Db 541 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAGGACCAACCATCGCAGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 711 GAGCCACACTTTCACATGTTGTAGCCAGAGAGGGGTGAGGACTCTTTCCTCGTGC 770
Db 661 GAGCCACACTTTCACATGTTGTAGCCAGAGAGGGGTGAGGACTCTTTCCTCGTGC 720
```


QY 771 GTAGCTTTTCATCTGTAAAGCTTCACTTAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 830
DB 721 GTAGCTTTTCATCTGTAAAGCTTCACTTAAGAGAGGAAACTTCTTGGTGTCTCAAGCAAG 780
QY 831 GAGATGGGCTCCAGCTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 890
DB 781 GAGATGGGCTCCAGCTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 840
QY 891 GGGAAAGCATCATCTCATGAAGGAAGAGATTTTGGCTGAAGAGGTGTACTCTCTCCA 950
DB 841 GGGAAAGCATCATCTCATGAAGGAAGAGATTTTGGCTGAAGAGGTGTACTCTCTCCA 900
QY 951 GATCTGGTGTCTTCTTGTGTGTAGATGTCCAGATGAAGCTTCATTCACCCATC 1010
DB 901 GATCTGGTGTCTTCTTGTGTGTAGATGTCCAGATGAAGCTTCATTCACCCATC 960
QY 1011 TGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGACAGATGCCCCCTGGCTTGGTG 1070
DB 961 TGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGACAGATGCCCCCTGGCTTGGTG 1020
QY 1071 GTTCACATGGCCCCAGCATCTGTGTGTGTGGACAGCAGGTACACAGAGTGGATGGAGG 1130
DB 1021 GTTCACATGGCCCCAGCATCTGTGTGTGTGGACAGCAGGTACACAGAGTGGATGGAGG 1080
QY 1131 TTTGGGCTTGACACCCAGCAGCTTGTCTTGAATGAGAACTGTGCTCAGTTCACAACTT 1190
DB 1081 TTTGGGCTTGACACCCAGCAGCTTGTCTTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
QY 1191 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTCTC 1250
DB 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTCTC 1200
QY 1251 ACCAGTTTCCGCTGTAAAGAGAGGCCCCACCTCAGTGTGCCATGGTTTCAGGGTGA 1310
DB 1201 ACCAGTTTCCGCTGTAAAGAGAGGCCCCACCTCAGTGTGCCATGGTTTCAGGGTGA 1260
QY 1311 TGCCTCTCAAGTACAGCTTCGCTCCAGAGGAGTGCACAGGATGCCATTATTACT 1370
DB 1261 TGCCTCTCAAGTACAGCTTCGCTCCAGAGGAGTGCACAGGATGCCATTATTACT 1320
QY 1371 TGCATCTTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGCAGAGCGTG 1430
DB 1321 TGCATCTTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGCAGAGCGTG 1380
QY 1431 CAGGAGTACAGGAGGAGTCCGACGAGCGGCCCGCCAGCAGGAGGAGTCAAGTAC 1490
DB 1381 CAGGAGTACAGGAGGAGTCCGACGAGCGGCCCGCCAGCAGGAGGAGTCAAGTAC 1440
QY 1491 CCAGAAATCATCTTCTTGGAAACAGGCTTCCATCCCATCCGATGAAATGTCAGT 1550
DB 1441 CCAGAAATCATCTTCTTGGAAACAGGCTTCCATCCCATCCGATGAAATGTCAGT 1500
QY 1551 GCCACTTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGACTGTGGTGGAGGCACA 1610
DB 1501 GCCACTTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGACTGTGGTGGAGGCACA 1560
QY 1611 TTTGGGACGTGTGCGCTCATTAACGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1670
DB 1561 TTTGGGACGTGTGCGCTCATTAACGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1620
QY 1671 GCTGTGTTTGTGTCCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1730
DB 1621 GCTGTGTTTGTGTCCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1731 CAGAGAGACGCGCTTGGCATTTTGGAAAGCCGCTTTCACCTTGTGCTGCTGCTGCTG 1790
DB 1681 CAGAGAGACGCGCTTGGCATTTTGGAAAGCCGCTTTCACCTTGTGCTGCTGCTGCTG 1740
QY 1791 CCCAACAGCTCAAGCCCTGGCTTCCAGCAGTACCAACACAGTGCAGGAGTCTCTGCAC 1850
DB 1741 CCCAACAGCTCAAGCCCTGGCTTCCAGCAGTACCAACACAGTGCAGGAGTCTCTGCAC 1800
QY 1851 CACATCAGTATGATTCCTGCCAAATGCTTTCAGGAAGGGGCTGAGATCTCCAGTCTGCA 1910

DB 1801 CACATCAGTATGATTCCTGCCAAATGCTTTCAGGAAGGGGCTGAGATCTCCAGTCTGCA 1860
QY 1911 GTGAAAGATTGATCAGTTTCGCTTGTGGAACATGTGATTTTGGAAAGAGTTTTCAGACCTGT 1970
DB 1861 GTGAAAGATTGATCAGTTTCGCTTGTGGAACATGTGATTTTGGAAAGAGTTTTCAGACCTGT 1920
QY 1971 CTGGTCCGCACTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACCTCTGGCTGGAAA 2030
DB 1921 CTGGTCCGCACTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACCTCTGGCTGGAAA 1980
QY 2031 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
DB 1981 GTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
QY 2091 ACCCTCTGATACATGAAGCCACCTTGAAGATGTTTGGAAAGAGAGTGGAAAAG 2150
DB 2041 ACCCTCTGATACATGAAGCCACCTTGAAGATGTTTGGAAAGAGAGTGGAAAAG 2100
QY 2151 ACACACAGCAAACTGCTCCAAAGCATCAGCTGGGATGCGGATGAACGCGAGTTCATT 2210
DB 2101 ACACACAGCAAACTGCTCCAAAGCATCAGCTGGGATGCGGATGAACGCGAGTTCATT 2160
QY 2211 ATGCTGAACCTTTCAGCCAGGCTATGCCAAGTCTCCCTTTCAGCCCAACTTCAGC 2270
DB 2161 ATGCTGAACCTTTCAGCCAGGCTATGCCAAGTCTCCCTTTCAGCCCAACTTCAGC 2220
QY 2271 GAGAAAGTGGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAACAATG 2330
DB 2221 GAGAAAGTGGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAACAATG 2280
QY 2331 CCCAAGCTGATTTCCCACTGAAAGCCTTGTGCTGCGCATCGAGGAGATGGAGAG 2390
DB 2281 CCCAAGCTGATTTCCCACTGAAAGCCTTGTGCTGCGCATCGAGGAGATGGAGAG 2340
QY 2391 CGCAGGAGAACGCGGAGCTCGGCGAGTGCAGGCGGCGCTTCTGTCAGGAGCTGGCA 2450
DB 2341 CGCAGGAGAACGCGGAGCTCGGCGAGTGCAGGCGGCGCTTCTGTCAGGAGCTGGCA 2400
QY 2451 GCGGCGCTTGGAGGATGGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCC 2510
DB 2401 GCGGCGCTTGGAGGATGGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCC 2460
QY 2511 AAGAAGGTTCAGAGCCCACTGA 2531
DB 2461 AAGAAGGTTCAGAGCCCACTGA 2481

RESULT 5

US-09-564-805-221
; Sequence 221, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR FILING DATE: 1998-11-06
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus

Db 2059 AGACACACAGACACACCTCCAGGCTATTAATGTGGGATCGGATGAATCGGAGTTC 2118
Qy 2208 ATTATGCTGAACACACTTCAGCCAGCCCTATGCCAAGGTCCCCCTCTTTCAGCCCCCACTTC 2267
Db 2119 ATCATGCTGAACACACTTCAGTCAGCGGTACGACGATCCCTCTTTCAGCCCTGACTTC 2178
Qy 2268 AGCGAAGAGTGGGAGTTCCTTTGACCAACATGAAGGTCTGTCTGGAGACTTTTCAACA 2327
Db 2179 AACGAGAAAGTTGGCATCGCTTTGACCAACATGAAGGTCTGTCTGGAGACTTTCCGACA 2238
Qy 2328 ATGCCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGACATCAGAGATGGAG 2387
Db 2239 GTGCCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGACATCAGAGATGGAG 2298
Qy 2388 GAGCGCAGGAGGAAGCGGAGGTGCGGCAGGTGCGGCGGCGCCCTCTGTCTCCAGGAGCTG 2447
Db 2299 GAACGAGGAGGAAGGAGGTACGCTGTGGGAGCAGCCCTCTGACC---CAGCAG 2355
Qy 2448 GCAGGCGGCTGGAGATGGGAGCTTCAGCAGAACGCGGCGCCACACAGAGAGCCACA 2506
Db 2356 GCAGACAGCCAGAGGACAGAGAACCCCAACAGAGAGCGGCGCCACACAGATGAACACA 2414

RESULT 6

US-09-564-805-28
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at

; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-564-805-28

Query Match 22.2%; Score 657.2; DB 4; Length 26664;
Best Local Similarity 99.5%; Pred. No. 9.6e-167;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2297 CATGAAGTCTGCTTTGGAGACTTTTCCAAACAATGCCAAGCTGATTTCCCCCTCTGAAAGC 2356
Db 25805 CTTCTAGGTCTGCTTTGGAGACTTTTCCAAACAATGCCAAGCTGATTTCCCCCTCTGAAAGC 25864
Qy 2357 CTTGTTTCTGCGACATCGAGGAGATGGAGAGCGCAGGAGAGAGCGGAGCTCGGCA 2416
Db 25865 CTTGTTTCTGCGACATCGAGGAGATGGAGAGCGCAGGAGAGAGCGGAGCTCGGCA 25924
Qy 2417 GGTGCGGCGGCGCTCTGTCTCAGGAGCTGGCAGCGGCGCTTGGAGGATGGGAGCCTCA 2476
Db 25925 GGTGCGGCGGCGCTCTGTCTCAGGAGCTGGCAGCGGCGCTTGGAGGATGGGAGCCTCA 25984
Qy 2477 GCAGAAGCGGCGCCACACAGAGGAGCCACAGCCCAAGAGGTCAGAGCCCACTGAAGATC 2536
Db 25985 GCAGAAGCGGCGCCACACAGAGGAGCCACAGCCCAAGAGGTCAGAGCCCACTGAAGATC 26044
Qy 2537 TGGGAGACCTGAACTCAGAGGCTGTGTCTTCTGCCCCCAGCCAGCCGCTATCTG 2596
Db 26045 TGGGAGACCTGAACTCAGAGGCTGTGTCTTCTGCCCCCAGCCAGCCGCTATCTG 26104
Qy 2597 CCCTCTTGTGTTAGAGCTGAAGAGCAGCGTCCCCCAGGAGGAGCTCAGGATAGGTG 2656
Db 26105 CCCTCTTGTGTTAGAGCTGAAGAGCAGCGTCCCCCAGGAGGAGCTCAGGATAGGTG 26164
Qy 2657 GTATGAGCTGTGCGAGGCTTGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTCA 2716
Db 26165 GTATGAGCTGTGCGAGGCTTGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTCA 26224
Qy 2717 GGACTGTGCTGGCAGCAGCGCGGCGGAGGCTGCCACAGGAGGAGGAGTGAAGTGA 2776
Db 26225 GGACTGTGCTGGCAGCAGCGCGGCGGAGGCTGCCACAGGAGGAGGAGTGAAGTGA 26284
Qy 2777 CTAATTTCAATTTCAAGGAGCTTTTAAAGAGCTTTGAAACAGACGCGGCGACCTTTCC 2836
Db 26285 CTAATTTCAATTTCAAGGAGCTTTTAAAGAGCTTTGAAACAGACGCGGCGACCTTTCC 26344
Qy 2837 TCTAATCCAGAAAGTATTTCCCTGCACACAGAGAGGAGTGAAGTGAAGTGAAGTGA 2896
Db 26345 TCTAATCCAGAAAGTATTTCCCTGCACACAGAGAGGAGTGAAGTGAAGTGAAGTGA 26404
Qy 2897 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTTCAGTGAATGAAGTGAAGTGA 2956
Db 26405 GTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTTTCAGTGAATGAAGTGAAGTGA 26464
Qy 2957 AA 2958
Db 26465 AA 26466

RESULT 7

US-09-564-805-27
; Sequence 27, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06

; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: exon 24
; NAME/KEY: polyA signal
; LOCATION: (636)..(641)
US-09-564-805-27

Query Match 22.1%; Score 655; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.2e-167;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2304 GTCTGCTTTGGAGACTTCCCAACATGCCCAGCTGATTCGCCCACTGAAAGCCCTGTTT 2363
DB 1 GTCTGCTTTGGAGACTTCCCAACATGCCCAGCTGATTCGCCCACTGAAAGCCCTGTTT 60
QY 2364 GCTGCGCATCGAGGAGATGAGGAGCGCAGGAGAGAGCGGAGCTCGCGCAGTGCGG 2423
DB 61 GCTGCGCATCGAGGAGATGAGGAGCGCAGGAGAGAGCGGAGCTCGCGCAGTGCGG 120
QY 2424 GCGGCCCTCTCTCCAGGAGCTGCGCAGCGGCTTGGAGGATGGGAGCCCTCAGCAGAG 2483
DB 121 GCGGCCCTCTCTCCAGGAGCTGCGCAGCGGCTTGGAGGATGGGAGCCCTCAGCAGAG 180
QY 2484 CCGGCCACACAGAGAGCCACAGGCGAAGAGTCAAGCCAGTGAAGATCTGGGAGA 2543
DB 181 CCGGCCACACAGAGAGCCACAGGCGAAGAGTCAAGCCAGTGAAGATCTGGGAGA 240
QY 2544 CCTCAACTCAGAAGCTGTGTCTCTCTGCCCCACGACGACCCCTATCTGCCCTCT 2603
DB 241 CCTCAACTCAGAAGCTGTGTCTCTCTGCCCCACGACGACCCCTATCTGCCCTCT 300
QY 2604 TGCTGTTAGAGCTGAAGAGCAGCGTCCGCCAGGAGGAGCTCAGGATAGTGTATGGA 2663
DB 301 TGCTGTTAGAGCTGAAGAGCAGCGTCCGCCAGGAGGAGCTCAGGATAGTGTATGGA 360
QY 2664 GCTGTGCGAGCTGGCTGCCACATAGACTAGTCTATAGATGCCCTCTTAGACTGG 2723
DB 361 GCTGTGCGAGCTTTGGGCTCCACATAGCACTAGTCTATAGATGCCCTCTTAGACTGG 420
QY 2724 TGCTGCGACAGCGCGGCGGAGGCTGCCACAGCGAAGCAAGCAGATGAATTT 2783
DB 421 TGCTGCGACAGCGCGGCGGAGGCTGCCACAGCGAAGCAAGCAGATGAATTT 480
QY 2784 CATTTCAAGGCAAGTTTTAAAGAGTCTTTGAAACAGAGCGGCGCACCTTTCTCTAATC 2843
DB 481 CATTTCAAGGCAAGTTTTAAAGAGTCTTTGAAACAGAGCGGCGCACCTTTCTCTAATC 540
QY 2844 CAGCAAGTATTCCTGTCACACAGAGCAAGAGTAAAGATCAGTGGTCTAAG 2903
DB 541 CAGCAAGTATTCCTGTCACACAGAGCAAGAGTAAAGATCAGTGGTCTAAG 600
QY 2904 TGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGTTGCAA 2958
DB 601 TGTCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTGTTGCAA 655

RESULT 8
US-09-564-805-210
; Sequence 210, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques

; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-564-805-210

Query Match 10.1%; Score 297.4; DB 4; Length 350;
Best Local Similarity 98.0%; Pred. No. 9.9e-71;
Matches 301; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGCGGGCTAGGTGACCGCGGCTTTCTCAGTTTGGTGAGAGCGGCGCATGTGGGCGC 60
DB 1 CGCGGGCTAGGTGACCGCGGCTTTCTCAGTTTGGTGAGAGCGGCGCATGTGGGCGC 60
QY 61 TTTGCTCGCTGCTGCGGTCGCGGCGGACGACCATGTCCGAGGAGCGACCATATCGC 120
DB 61 TTTGCTCGCTGCTGCGGTCGCGGCGGAGCGACCATGTCCGAGGAGCGACCATATCGC 120
QY 121 AGGCACCGCCCGCGCGAGCGCGCAAGAGACCCGCTGCGGCGACCTGCGCACGCGAG 180
DB 121 AGGCACCGCCCGCGCGAGCGCGCAAGAGACCCGCTGCGGCGACCTGCGCACGCGAG 180
QY 181 AGAAGCGGACCGTCCGGGTGCTCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 181 AGAAGCGGACCGTCCGGGTGCTCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 CAGCGGCTAGCCGGGACTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 CAGCGGCTAGCCGGGACTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 TCTTCAA 307
DB 301 TCAACGA 307

RESULT 9
US-09-564-805-4
; Sequence 4, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4

;
; LENGTH: 295
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(295)
; OTHER INFORMATION: exon 1
US-09-564-805-4

Query Match 10.0%; Score 295; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGGCTAGTACCGCGGCTTCTCAGTTTGGTGGAGACGGGCCATGTGGGGCC 60
Db 1 CGCGGGCTAGTACCGCGGCTTCTCAGTTTGGTGGAGACGGGGCATGTGGGGCC 60
Qy 61 TTGTGCTGCTGCTGCGGTCCGCGCGGAGCGACCATGTGCGAGGAGCCACCATATCGC 120
Db 61 TTGTGCTGCTGCTGCGGTCCGCGCGGAGCGACCATGTGCGAGGAGCCACCATATCGC 120
Qy 121 AGCACCCCGCCGCGAGCGCGCGCAAGACCCGCTGCGGACCTGCGCAGCGCGAG 180
Db 121 AGCACCCCGCCGCGAGCGCGCGCAAGACCCGCTGCGGACCTGCGCAGCGCGAG 180
Qy 181 AGAAGCGCGAGCGTGGGGTGTCCGGCGGCCCAACACCGTGTACCTGCAAGTGGTGG 240
Db 181 AGAAGCGCGAGCGTGGGGTGTCCGGCGGCCCAACACCGTGTACCTGCAAGTGGTGG 240
Qy 241 CAGCGGTTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCCGAGTTCAACCG 295
Db 241 CAGCGGTTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCCGAGTTCAACCG 295

RESULT 10

US-09-328-111-315
; Sequence 315, Application US/09328111
; Patent No. 6262333

; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-111-315

Query Match 8.0%; Score 237; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 226 ACCTGAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCGG 285
Db 1 ACCTGAGGTGGTGGCAGCGGTAGCGGACTCGGCGCGCGCTCTACGCTCTCTCGG 60

Qy 286 AGTTCAACCGGTATCTTTCAACTGTGTGAGAGGCGTTTCAAGACTCATGTGAGGAGCACA 345
Db 61 AGTTCAACCGGTATCTTTCAACTGTGTGAGAGGCGTTTCAAGACTCATGTGAGGAGCACA 120
Qy 346 AGTTAAAGGTTGCTCGCCTGGACACATATTCCTTACACAGATGCACTGGTCTTAATGTTG 405
Db 121 AGTTAAAGGTTGCTCGCCTGGACACATATTCCTTACACAGATGCACTGGTCTTAATGTTG 180
Qy 406 GGGGCTTAAGTGAATGATTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAC 462
Db 181 GGGGCTTAAGTGAATGATTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAC 237

RESULT 11

US-09-564-805-26
; Sequence 26, Application US/09564805
; Patent No. 6333403

; GENERAL INFORMATION:
; APPLICANT: Tavtligian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(145)
; OTHER INFORMATION: exon 23
US-09-564-805-26

Query Match 4.9%; Score 145; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.7e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2159 CACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGAGTTTCATTATCTGAA 2218
Db 1 CACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGAGTTTCATTATCTGAA 60
Qy 2219 CCATTTCAGCCAGCGCTATGCCAAGTCCCTCTTACGCCCCCACTTCAGCGAGAAAGT 2278
Db 61 CCATTTCAGCCAGCGCTATGCCAAGTCCCTCTTACGCCCCCACTTCAGCGAGAAAGT 120
Qy 2279 GGGAGTTCCTTTGACCACATGAAG 2303
Db 121 GGGAGTTCCTTTGACCACATGAAG 145

RESULT 12

US-09-564-805-16
; Sequence 16, Application US/09564805
; Patent No. 6333403

; GENERAL INFORMATION:
; APPLICANT: Tavtligian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility

;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/564,805
;; CURRENT FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 139
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(139)
;; OTHER INFORMATION: exon 13
US-09-564-805-16

Query Match 4.7%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.1e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1130 GTTTGGGCTGACACCCAGCACTTGGTCTGAATGAGAACTGTGCTCAGTTCAACAACCT 1189
Db 1 GTTTGGGCTGACACCCAGCACTTGGTCTGAATGAGAACTGTGCTCAGTTCAACAACCT 60
Qy 1190 TCGAGCCACAAAGATTCAAAACCCAGCACTCAACCTCATCCACCGGACATCTTCCCCCTGCT 1249
Db 61 TCGAGCCACAAAGATTCAAAACCCAGCACTCAACCTCATCCACCGGACATCTTCCCCCTGCT 120
Qy 1250 CACCAGTTTCGCTGTAAG 1268
Db 121 CACCAGTTTCGCTGTAAG 139

RESULT 13
US-09-564-805-20
;; Sequence 20, Application US/09564805
;; Patent No. 6333403
;; GENERAL INFORMATION:
;; APPLICANT: Tavtigian, Sean V.
;; APPLICANT: Teng, David H.F.
;; APPLICANT: Simard, Jacques
;; APPLICANT: Rommens, Johanna M.
;; APPLICANT: Myriad Genetics, Inc.
;; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/564,805
;; CURRENT FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 20
;; LENGTH: 139
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(139)
;; OTHER INFORMATION: exon 17
US-09-564-805-20

Query Match 4.7%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.1e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1571 CCCCAGACGTCCTGCTACTGAGCACTGGTGAGGGCACATTGGGCACGCTGTGCCGTCA 1630

Db 1 CCCCAGACGTCCTGCTACTGAGCACTGGTGAGGGCACATTGGGCAGCTGTGCCGTCA 60
Qy 1631 TTACGGAGACAGGTGACAGGGTCTGGGACACCTCGCTGTGTGTTGTGTCACCT 1690
Db 61 TTACGGAGACAGGTGACAGGGTCTGGGACACCTCGCTGTGTGTTGTGTCACCT 120
Qy 1691 GCAGGAGATCACCACAG 1709
Db 121 GCAGGAGATCACCACAG 139

RESULT 14
US-09-564-805-24
;; Sequence 24, Application US/09564805
;; Patent No. 6333403
;; GENERAL INFORMATION:
;; APPLICANT: Tavtigian, Sean V.
;; APPLICANT: Teng, David H.F.
;; APPLICANT: Simard, Jacques
;; APPLICANT: Rommens, Johanna M.
;; APPLICANT: Myriad Genetics, Inc.
;; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/564,805
;; CURRENT FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 24
;; LENGTH: 121
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(121)
;; OTHER INFORMATION: exon 21
US-09-564-805-24

Query Match 4.1%; Score 121; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1959 TTTCAGACCTCTCTGGTCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACC 2018
Db 1 TTTCAGACCTCTCTGGTCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACC 60
Qy 2019 TCTGGTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATG 2078
Db 61 TCTGGTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATG 120
Qy 2079 G 2079
Db 121 G 121

RESULT 15
US-09-564-805-10
;; Sequence 10, Application US/09564805
;; Patent No. 6333403
;; GENERAL INFORMATION:
;; APPLICANT: Tavtigian, Sean V.
;; APPLICANT: Teng, David H.F.
;; APPLICANT: Simard, Jacques
;; APPLICANT: Rommens, Johanna M.
;; APPLICANT: Myriad Genetics, Inc.
;; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120)
; OTHER INFORMATION: exon 7
US-09-564-805-10

Query Match 4.1%; Score 120; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.8e-23;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 610 GTGAACAGAGGAGGGGAAGACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGC 669
|||
Db 1 GTGAACAGAGGAGGGGAAGACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGC 60
|||
Oy 670 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 729
|||
Db 61 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 120
|||

Search completed: May 18, 2003, 00:09:50
Job time : 249 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2003, 00:09:53 ; Search time 359 Seconds
(without alignments)
10628.516 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcggcgtagtgaccggc.....ataaagattgattgtgcaa 2958

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 810007 seqs, 644960991 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2958	100.0	2958	9	US-09-988-626-3
2	2958	100.0	2958	9	US-09-988-687-3
3	2874.4	97.2	2908	9	US-09-988-626-223
4	2874.4	97.2	2908	9	US-09-988-687-223
5	2819.6	95.3	2892	9	US-09-988-626-225
6	2819.6	95.3	2892	9	US-09-988-687-225
7	2481	83.9	2481	9	US-09-988-626-1
8	1645.6	55.6	2470	9	US-09-988-687-1
9	1645.6	55.6	2470	9	US-09-988-626-221
10	1645.6	55.6	2470	9	US-09-988-687-221
11	734.8	24.8	783	10	US-09-833-381-2039
12	657.2	22.2	26664	9	US-09-988-626-28
13	657.2	22.2	26664	9	US-09-988-687-28
14	655	22.1	655	9	US-09-988-626-27
15	655	22.1	655	9	US-09-988-687-27
16	470.4	15.9	536	10	US-09-833-381-2038
17	432.8	14.6	554	9	US-09-918-995-8996
18	297.4	10.1	350	9	US-09-988-626-210
19	297.4	10.1	350	9	US-09-988-687-210

Sequence 4, Appli
Sequence 4, Appli
Sequence 315, App
Sequence 26, Appl
Sequence 26, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 212, App
Sequence 212, App
Sequence 14, Appl
Sequence 14, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 15, Appl

20 295 10.0 295 9 US-09-988-626-4
21 295 10.0 295 9 US-09-988-687-4
22 237 8.0 238 10 US-09-879-536-315
23 145 4.9 145 9 US-09-988-626-26
24 145 4.9 145 9 US-09-988-687-26
25 139 4.7 139 9 US-09-988-626-16
26 139 4.7 139 9 US-09-988-626-20
27 139 4.7 139 9 US-09-988-687-16
28 139 4.7 139 9 US-09-988-687-20
29 121 4.1 121 9 US-09-988-626-24
30 121 4.1 121 9 US-09-988-687-24
31 120 4.1 120 9 US-09-988-626-10
32 120 4.1 120 9 US-09-988-687-10
33 119 4.0 119 9 US-09-988-626-18
34 119 4.0 119 9 US-09-988-687-18
35 118.4 4.0 326 9 US-09-988-626-212
36 118.4 4.0 326 9 US-09-988-687-212
37 113 3.8 113 9 US-09-988-626-14
38 113 3.8 113 9 US-09-988-687-14
39 110 3.7 110 9 US-09-988-626-22
40 110 3.7 110 9 US-09-988-687-22
41 100 3.4 100 9 US-09-988-626-23
42 100 3.4 100 9 US-09-988-687-23
43 97 3.3 97 9 US-09-988-626-19
44 97 3.3 97 9 US-09-988-687-19
45 96 3.2 96 9 US-09-988-626-15

ALIGNMENTS

RESULT 1

US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match 100.0%; Score 2958; DB 9; Length 2958;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGCGGCGTAGTGACCGCGGCTTCTCAGTTTGGTGACGCGGCGCATGTGGCGC 60

Db 1 CCGCGGCGTAGTGACCGCGGCTTCTCAGTTTGGTGACGCGGCGCATGTGGCGC 60

Qy 61 TTGTGCTGCTGCTGCGGTCGCGGCGGCGGACGACCATGTGCGAGGACGACCATATCGC 120

Db 61 |||||TTTGTCTGCTGCGTCCGCGCGGAGCGACCATGTGCGAGGAGCGACCATATGCG 120
Qy 121 AGGCAACCGCCCGCGCGAGCGCGCGAGAACGCCGCTGCGGCACTTGGCGCAGCGAG 180
Db 121 AGGCAACCGCCCGCGCGAGCGCGCGAGAACGCCGCTGCGGCACTTGGCGCAGCGAG 180
Qy 181 AGAAGCGGGACCGTTCGGGGTCTCCGGGGCGCCCAACACCGTGTACTTGCAGGTGGTGG 240
Db 181 AGAAGCGGGACCGTTCGGGGTCTCCGGGGCGCCCAACACCGTGTACTTGCAGGTGGTGG 240
Qy 241 CAGCGGGTAGCGGACTTCGGCGCGCGGCTCTACGTCTTCTCCGAGTTCAACCGGTATC 300
Db 241 CAGCGGGTAGCGGACTTCGGCGCGCGGCTCTACGTCTTCTCCGAGTTCAACCGGTATC 300
Qy 301 TCTTCAACTGTGGAGAGCGTTACAGACTCATCCAGGAGCACAAAGTTAAAGTTGCTC 360
Db 301 TCTTCAACTGTGGAGAGCGTTACAGACTCATCCAGGAGCACAAAGTTAAAGTTGCTC 360
Qy 361 GCCTGGACAACATATCTGACACCAATGCATCTGCTTAATGTTGGGGGCTTAAAGTGAA 420
Db 361 GCCTGGACAACATATCTGACACCAATGCATCTGCTTAATGTTGGGGGCTTAAAGTGAA 420
Qy 421 TGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGACCTCCACAAC 480
Db 421 TGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTACTTTCTGGACCTCCACAAC 480
Qy 481 TGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGAATAGAACTGG 540
Db 481 TGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGAATAGAACTGG 540
Qy 541 CTGTGGGGCCCCACTCTGCCCCAGATACGAGGATGAACCATGACAGTTTACAGATCC 600
Db 541 CTGTGGGGCCCCACTCTGCCCCAGATACGAGGATGAACCATGACAGTTTACAGATCC 600
Qy 601 CCATACACAGTGAAACAGAGGAGGGAAGACCAACCATGCGAGAGTCAGAAAGGCGCTC 660
Db 601 CCATACACAGTGAAACAGAGGAGGGAAGACCAACCATGCGAGAGTCAGAAAGGCGCTC 660
Qy 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Db 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAATGAGCCACACC 720
Qy 721 TTCACATGGTGTAGCAGAGAGAGGGGTACAGGACTTCCCTGCTGTAGCTTTCA 780
Db 721 TTCACATGGTGTAGCAGAGAGAGGGGTACAGGACTTCCCTGCTGTAGCTTTCA 780
Qy 781 TCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAAGCAAGAGAGTGGGCC 840
Db 781 TCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGTCTCAAAGCAAGAGAGTGGGCC 840
Qy 841 TCCCAGTTGGGACAGCTGCCATCGCTCCCATCACTGCTGTCAAGACGGGAAAAGCA 900
Db 841 TCCCAGTTGGGACAGCTGCCATCGCTCCCATCACTGCTGTCAAGACGGGAAAAGCA 900
Qy 901 TCACATCAAGAGAGAGATTTTGGCTGAAGACTGTGTACTCTCCAGATCCGTGGT 960
Db 901 TCACATCAAGAGAGAGATTTTGGCTGAAGACTGTGTACTCTCCAGATCCGTGGT 960
Qy 961 CTGCTTTTGTGGTGTAGAAATCTCCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAATG 1020
Db 961 CTGCTTTTGTGGTGTAGAAATCTCCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAATG 1020
Qy 1021 CCACCTTTCAAGGTACCAAGAAAGGAGATGCCCGTGGCTTGTGGTTTCAATGG 1080
Db 1021 CCACCTTTCAAGGTACCAAGAAAGGAGATGCCCGTGGCTTGTGGTTTCAATGG 1080
Qy 1081 CCCAGCATCTGTGCTGTGGACAGAGGTACACAGTGGATGGAGAGTTTGGGCTG 1140
Db 1081 CCCAGCATCTGTGCTGTGGACAGAGGTACACAGTGGATGGAGAGTTTGGGCTG 1140
Qy 1141 ACACCCAGCAGCTTGTCTGTAATGAAACTGTGCTCTCAGTTTCAACAACCTTCGACCCACA 1200

Db 1141 ACACCCAGCAGCTTGTCTGTAATGAAACTGTGCTCTCAGTTTCAACAACCTTCGACGCCACA 1200
Qy 1201 AGATTCAAAACCCAGCTCAACTCATCCACCGGAGCATCTTCCCCTCTGCTACACGATTTTC 1260
Db 1201 AGATTCAAAACCCAGCTCAACTCATCCACCGGAGCATCTTCCCCTCTGCTACACGATTTTC 1260
Qy 1261 GCTCTAAGAGGAGGGCCCAACCTCAGTGTGCCCATGGTTTCAGGGTGAATGCTCTCTCA 1320
Db 1261 GCTCTAAGAGGAGGGCCCAACCTCAGTGTGCCCATGGTTTCAGGGTGAATGCTCTCTCA 1320
Qy 1321 AGTACAGCTCCGTCCAGGAGGAGTGGCAGAGGAGTGCATTATTACTTGCATCTCTG 1380
Db 1321 AGTACAGCTCCGTCCAGGAGGAGTGGCAGAGGAGTGCATTATTACTTGCATCTCTG 1380
Qy 1381 AGGAATTTCAATAGTTAGGGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAGAGTACA 1440
Db 1381 AGGAATTTCAATAGTTAGGGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAGAGTACA 1440
Qy 1441 GGAGGAGTGCAGGAGCGGCCAGCCAGCCAGCAGAGAAAGAGTCACTCCACAGAAATCA 1500
Db 1441 GGAGGAGTGCAGGAGCGGCCAGCCAGCCAGCAGAGAAAGAGTCACTCCACAGAAATCA 1500
Qy 1501 TCTTCTTTGAAACAGGGTCTGCCATCCGATGAAGATTTCGAAATGTTCAGTGCACACTTG 1560
Db 1501 TCTTCTTTGAAACAGGGTCTGCCATCCGATGAAGATTTCGAAATGTTCAGTGCACACTTG 1560
Qy 1561 TCAACATAAGCCCCGACACAGCTCTCTGCTACTGGAATGTGGTGAAGGACACTTTGGGAGC 1620
Db 1561 TCAACATAAGCCCCGACACAGCTCTCTGCTACTGGAATGTGGTGAAGGACACTTTGGGAGC 1620
Qy 1621 TGTGCCCTCATTTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGGCTGTGTGTTG 1680
Db 1621 TGTGCCCTCATTTACGGAGACAGGTGACAGGGTCTTGGGACCCCTGGCTGTGTGTTG 1680
Qy 1681 TGTCCCACCTGCAGCAGATCACACACGGGCTTGCACAGTATCTTGTGCAGAGAGAAC 1740
Db 1681 TGTCCCACCTGCAGCAGATCACACACGGGCTTGCACAGTATCTTGTGCAGAGAGAAC 1740
Qy 1741 GCGCTTTGGGATCTTTGGGAAAGCGCTTCACTTGTGTGGTGGTGGTGGTGGTGGTGGTGG 1800
Db 1741 GCGCTTTGGGATCTTTGGGAAAGCGCTTCACTTGTGTGGTGGTGGTGGTGGTGGTGGTGG 1800
Qy 1801 TCAAGCCTCGCTCCAGCAGTACCAACACAGTCCAGAGGTCCTGCACACATCAGTA 1860
Db 1801 TCAAGCCTCGCTCCAGCAGTACCAACACAGTCCAGAGGTCCTGCACACATCAGTA 1860
Qy 1861 TGATTCCTGCAAAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTTCAGTGGAAAGAT 1920
Db 1861 TGATTCCTGCAAAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTTCAGTGGAAAGAT 1920
Qy 1921 TGATCAGTTCTGCTTCCGAACATGTGATTTGGAAGAGTTTTCAGACCTGTCTGTGGGGC 1980
Db 1921 TGATCAGTTCTGCTTCCGAACATGTGATTTGGAAGAGTTTTCAGACCTGTCTGTGGGGC 1980
Qy 1981 ACTCAAGCATGGTTTGGCTGTGCGTGTGCACACTCTGGCTGGAAGTGTCTATT 2040
Db 1981 ACTCAAGCATGGTTTGGCTGTGCGTGTGCACACTCTGGCTGGAAGTGTCTATT 2040
Qy 2041 CCGGGGACACCATGCTCCCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCCACCTCTCTGA 2100
Db 2041 CCGGGGACACCATGCTCCCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCCACCTCTCTGA 2100
Qy 2101 TACATGAAGCCACCTTGAAGATGGTTTGAAGGAAGCAGTGGAAAGACACACAGCA 2160
Db 2101 TACATGAAGCCACCTTGAAGATGGTTTGAAGGAAGCAGTGGAAAGACACACAGCA 2160
Qy 2161 CAACGCTCCCAAGCCATCAGCGTGGGATGGGATGAACGGAGTTCATTATGCTGAACC 2220
Db 2161 CAACGCTCCCAAGCCATCAGCGTGGGATGGGATGAACGGAGTTCATTATGCTGAACC 2220
Qy 2221 ACTTCAGCCAGCGCTATGCCAAGGTCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280
Db 2221 ACTTCAGCCAGCGCTATGCCAAGGTCCTCTTTCAGCCCCCACTTCAGCGAGAAAGTGG 2280

Qy	2281	GAGTTGCTTTGACACACATGAAGGCTGCTTTTGGAGACTTTTCCAACAATGCCCAAGCTGA	2340
Db	2281	GAGTTGCTTTTGGACACATGAAGGCTGCTTTTGGAGACTTTTCCAACAATGCCCAAGCTGA	2340
Qy	2341	TTCCCCACACTGAAGCCCTGTTTGTGGCGCATCGAGAGATCGAGGAGCGCAGGGAGA	2400
Db	2341	TTCCCCCACTGAAGAACCCTGTTTGTGGCGCATCGAGAGATCGAGGAGCGCAGGGAGA	2400
Qy	2401	AGCGGGAGCTGCGGCAGGTCGCGGGCGCCCTCTCTGCCAGGGAGCTGGCAGCGCGCTGG	2460
Db	2401	AGCGGAGCTGCGGCAGGTCGCGGGCGCCCTCTCTGCCAGGGAGCTGGCAGCGCGCTGG	2460
Qy	2461	AGGATGGGAGCCTCAGCAGAGACGGGGCCCAACACAGAGAGAGCCACAGGCCAAGAAGTCA	2520
Db	2461	AGGATGGGGAGCCTCAGCAGAGAAGCGGGCCCAACACAGAGAGAGCCACAGGCCAAGAAGTCA	2520
Qy	2521	GAGCCCACTCAAGATCTGGGAGACCTGAACCTCAGAAGGCTGTGTGTTCTTCGCCCCACG	2580
Db	2521	GAGCCCACTGAAGATCTGGGAGACCCTGAACCTCAGAAGGCTGTGTGTTCTTCGCCCCACG	2580
Qy	2581	CACGACCCGCTATCTGCCCTCTCTGTGTGTAGAACTGAAAGACACGGTCCCCACGAGG	2640
Db	2581	CACGACCCGCTATCTGCCCTCTCTGTGTGTAGAACTGAAAGACACGGTCCCCACGAGG	2640
Qy	2641	CAGCTCAGGATAGGTGGTATGAGCTGTGCCAGGCTTGGGCTCCCATAGACACTAGT	2700
Db	2641	CAGCTCAGGATAGGTGGTATGAGAGCTGTGCCAGGCTTGGGCTCCCATAGACACTAGT	2700
Qy	2701	CTATAGATGCTCTTAGGACTGTGGCTGGCACACGCGGGCCAGGAGGCTGCCACACG	2760
Db	2701	CTATAGATGCTCTTAGGACTGTGGCTGGCACACGCGGGCCAGGAGGCTGCCACACG	2760
Qy	2761	GAAGCAAGCAGATGAACCTAAATTCATTTCAAGGACAGTCTTTAAAGAACTCTTGAAACAG	2820
Db	2761	GAAGCAAGCAGATGAACCTAAATTCATTTCAAGGACAGTCTTTAAAGAACTCTTGAAACAG	2820
Qy	2821	ACGGCGCACCTTCTCTAATCCAGCAAGATGATTCCCTGCACACACGAGACAAAGCAGA	2880
Db	2821	ACGGCGCACCTTCTCTAATCCAGCAAGATGATTCCCTGCACACACGAGACAAAGCAGA	2880
Qy	2881	GTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTTAAACGAAAATAGTATTTTCAGCTCAA	2940
Db	2881	GTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTTAAACGAAAATAGTATTTTCAGCTCAA	2940
Qy	2941	TAAAGATTGAGTTTGCAG	2958
Db	2941	TAAAGATTGAGTTTGCAG	2958

RESULT. 2.

```

US-09-398-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Tavtgian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
;
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
;
; FILE REFERENCE: 2318-259
;
; CURRENT APPLICATION NUMBER: US/09/988,687
;
; CURRENT FILING DATE: 2001-11-20
;
; PRIOR APPLICATION NUMBER: 09/564,805
;
; PRIOR FILING DATE: 2000-05-05
;
; PRIOR APPLICATION NUMBER: US 60/107,468
;
; PRIOR FILING DATE: 1998-11-06
;
; PRIOR APPLICATION NUMBER: 09/434,382
;
; PRIOR FILING DATE: 1999-11-05
;
; NUMBER OF SEQ ID NOS: 240
;
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-687-3

Query Match      100.0%; Score 2958; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGCGTAGGTGACCGGGCGCTTCTCAGTTTGTGGAGACGGCGCATGTGGGCGC 60
Db 1 CGCGGCGTAGGTGACCGGGCGCTTCTCAGTTTGTGGAGACGGCGCATGTGGGCGC 60

Qy 61 TTTGCTCGTCTGCTGGGTCCGGGCGGAGCGACCATGTGCGAGGACGACCATATCGC 120
Db 61 TTTGCTCGTCTGCTGGGTCCGGGCGGAGCGACCATGTGCGAGGACGACCATATCGC 120

Qy 121 AGGCACCGCCGCGCGAGCGGGCGGCCAAGGACCCGCTCGGCACCTGCGCACGCGAG 180
Db 121 AGGCACCGCCGCGCGAGCGGGCGGCCAAGGACCCGCTCGGCACCTGCGCACGCGAG 180

Qy 181 AGAAGCGCGGACCGTGGGGTGCTCCGGCGGCCCAAAACACCGTGTACCTGCAGGTGTGG 240
Db 181 AGAAGCGCGGACCGTGGGGTGCTCCGGCGGCCCAAAACACCGTGTACCTGCAGGTGTGG 240

Qy 241 CAGCGGTAGCCGGGACTCGGGCGCGCGCTCTACGTTCTTCGAGTTCAACCGGTATC 300
Db 241 CAGCGGTAGCCGGGACTCGGGCGCGCGCTCTACGTTCTTCGAGTTCAACCGGTATC 300

Qy 301 TCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGACGAGCACAAGTTAAAGGTTGCTC 360
Db 301 TCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGAGAGCACAAGTTAAAGGTTGCTC 360

Qy 361 GCCTGGACAAACATATTCCTGACAGAAATGCACTGGTCTAATGTTGGGGGCTTAACTGGAA 420
Db 361 GCCTGGACAAACATATTCCTGACAGAAATGCACTGGTCTAATGTTGGGGGCTTAACTGGAA 420

Qy 421 TGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGACCTCCACAAC 480
Db 421 TGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGACCTCCACAAC 480

Qy 481 TGGAAATAACTCGAGCAATCAAAATATTTCTGGTCCATTGAAAGGAATAGAACTGG 540
Db 481 TGGAAATAACTCGAGCAATCAAAATATTTCTGGTCCATTGAAAGGAATAGAACTGG 540

Qy 541 CTGTGCGGCCCACTCTCCCCAGAAATACGAGGATGAAACCATGACAGTTTACCAAGATCC 600
Db 541 CTGTGCGGCCCACTCTCCCCAGAAATACGAGGATGAAACCATGACAGTTTACCAAGATCC 600

Qy 601 CCATACACAGTGAACAGAGGGGAAAGCACCAACCATGCGAGAGTCCGAAAGGCGCTC 660
Db 601 CCATACACAGTGAACAGAGGGGAAAGCACCAACCATGCGAGAGTCCGAAAGGCGCTC 660

Qy 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAATGAAATGAGCCACACC 720
Db 661 TCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAATGAAATGAGCCACACC 720

Qy 721 TTCCACATGGTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGTCGTAGCTTTCA 780
Db 721 TTCCACATGGTGTAGCCAGAGAGAGGGGTCAAGGACTCTTCCCTGTCGTAGCTTTCA 780

Qy 781 TCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAAGCAAGGAGATGGGCC 840
Db 781 TCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAAGCAAGGAGATGGGCC 840

Qy 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGACGGGAAAGCA 900
Db 841 TCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGGACGGGAAAGCA 900

```

Qy	901	TCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTG	960
Db	901	TCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTG	960
Qy	961	CTGCTTTTTGTGGTGTAGANTGTCAGATGAAGCTTCAATTCAAACCCATCTGTGAGATG	1020
Db	961	CTGCTTTTTGTGGTGTAGAATGTCCAAGATGAAGAGCTTCAATTCAAACCCATCTGTGAGAA	1020
Qy	1021	CCACCTTTACAGAGTACAAGGAAAGCAGATGCCCGTGGCCTTGGTGGTTCCACATGG	1080
Db	1021	CCACCTTTACAGAGTACCAGGAAAGCAGATGCCCGTGGCCTTGGTGGTTCCACATGG	1080
Qy	1081	CCCCAGCATCTGTCTTGTGGACAGCAGGTACACAGTCAGGTGAGAGGTTTGGGCCTG	1140
Db	1081	CCCCAGCATCTGTCTTGTGGACAGCAGGTACACAGTCAGGTGAGAGGTTTGGGCCTG	1140
Qy	1141	ACACCCAGCACTTGGTCTGTAATGAAACTGTGCTCAGTTCCAGAACCTTCGAGCCACA	1200
Db	1141	ACACCCAGCACTTGGTCTGTAATGAAACTGTGCTCAGTTCCAGAACCTTCGAGCCACA	1200
Qy	1201	AGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCTGCTCACAGTTTCC	1260
Db	1201	AGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCTGCTCACAGTTTCC	1260
Qy	1261	GCTGTAGAAGAGGAGGCCCCACCTCAGTGTGCCATGGTTCAAGGTGAATGCCCTCTCA	1320
Db	1261	GCTGTAGAAGAGGAGGCCCCACCTCAGTGTGCCATGGTTCAAGGTGAATGCCCTCTCA	1320
Qy	1321	AGTACAGCTCCGTCCACGAGGAGTGCCAGAGGATGCCATATTAATTCTTGCATCCTG	1380
Db	1321	AGTACAGCTCCGTCCACGAGGAGTGCCAGAGGATGCCATATTAATTCTTGCATCCTG	1380
Qy	1381	AGAAATTCATAGTTGAGGGCGTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAAGTAGA	1440
Db	1381	AGAAATTCATAGTTGAGGGCGTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAAGTAGA	1440
Qy	1441	GGAGGAGTGCAGAGACGGGCCAGCCCCAGCAGAGAAAGATCAGTACCAGAAATCA	1500
Db	1441	GGAGGAGTGCAGAGACGGGCCAGCCCCAGCAGAGAAAGATCAGTACCAGAAATCA	1500
Qy	1501	TC TTCTCTTGAAACAGGCTGTGCCATCCCGATCAAGATTGCAAAATGTTCAGTSCCACCTG	1560
Db	1501	TC TTCTCTTGAAACAGGCTGTGCCATCCCGATCAAGATTGCAAAATGTTCAGTSCCACCTG	1560
Qy	1561	TCAACATAAGCCCCCGACACGTCTCTGCTACTGGACTGTGTTGAGGGCACATTTGGGCAGC	1620
Db	1561	TCAACATAAGCCCCCGACACGTCTCTGCTACTGGACTGTGTTGAGGGCACATTTGGGCAGC	1620
Qy	1621	TGTGCGGTCAATTACGAGACCGAGTGGAAGGGTCTGGGCACCCCTGGCTGCTGTGTTG	1680
Db	1621	TGTGCGGTCAATTACGAGACCGAGTGGAAGGGTCTGGGCACCCCTGGCTGCTGTGTTG	1680
Qy	1681	TGTCCCACTGCAACGAGATCACCAACGGGCTTCCCAAGTATCTTGTCTGCAGAGAGAAC	1740
Db	1681	TGTCCCACTGCAACGAGATCACCAACGGGCTTCCCAAGTATCTTGTCTGCAGAGAGAAC	1740
Qy	1741	GC GCTTGGCATCTTTGGAAAAGCGCTTCAACCTTTGCTGGTGGTGTGCCCCAACCGC	1800
Db	1741	GC GCTTGGCATCTTTGGAAAAGCGCTTCAACCTTTGCTGGTGGTGTGCCCCAACCGC	1800
Qy	1801	TCAAAGCTGGCTCCAGCAGTACCAACACGAGTGCCAGAGGTCTCTGCACCATCAGTA	1860
Db	1801	TCAAAGCTGGCTCCAGCAGTACCAACACGAGTGCCAGAGGTCTCTGCACCATCAGTA	1860
Qy	1861	TGATTCTCTGCAAAATGCCITTCAGGAAGGGCTGAGATCTCCAGTCTCTGCAGTGGAAAGAT	1920
Db	1861	TGATTCTCTGCAAAATGCCITTCAGGAAGGGCTGAGATCTCCAGTCTCTGCAGTGGAAAGAT	1920
Qy	1921	TGATCAGTTCGCTTGTGGCAACATGTGATTTGGAAAGAGTTTCAGACCTGTCTGTGTCGGC	1980
Db	1921	TGATCAGTTCGCTTGTGGCAACATGTGATTTGGAAAGAGTTTCAGACCTGTCTGTGTCGGC	1980

Qy	1981	ACTCAAGCATGCGTTTGGCTGTGGCTGGTGCA	CACCTCTGGCTGGAAAGTGCTCTATT	2041
Db	1981	ACTCAAGCATGCGTTTGGCTGTGGCTGGTGCA	CACCTCTGGCTGGAAAGTGCTCTATT	2040
Qy	2041	CCGGGACACATGCGCTCGAGGCTCTGGTCCGAT	CGGGAAAGATGCGCACCCCTCTCTGA	2100
Db	2041	CCGGGACACATGCGCTCGAGGCTCTGGTCCGAT	CGGGAAAGATGCGCACCCCTCTCTGA	2100
Qy	2101	TACATGAAGCCACCTCGGAAGATGGTTTGGAA	GGAAGCAAGCAGTGGAAAAAGACACACAGCA	2160
Db	2101	TACATGAAGCCACCTCGGAAGATGGTTTGGAA	GGAAGCAAGCAGTGGAAAAAGACACACAGCA	2160
Qy	2161	CAACGTCCTCCAAAGCCATACGCTGGGGATG	CGCGAGTTCATTATGCTGAACC	2220
Db	2161	CAACGTCCTCCAAAGCCATACGCTGGGGATG	CGCGAGTTCATTATGCTGAACC	2220
Qy	2221	ACTTCAGCCAGCGCTATGCCAAAGTCCCTCT	TCCTCAGCCCCCACTTCAGCGAAGAGTGG	2280
Db	2221	ACTTCAGCCAGCGCTATGCCAAAGTCCCTCT	TCCTCAGCCCCCACTTCAGCGAAGAGTGG	2280
Qy	2281	GAGTTGCTTTTGACCAATCAAGAGTCTGCTTT	TGGAGACTTTCACCAATGCCCAAGCTGA	2340
Db	2281	GAGTTGCTTTTGACCAATCAAGAGTCTGCTTT	TGGAGACTTTCACCAATGCCCAAGCTGA	2340
Qy	2341	TTCCCCACATGAAGCCCTGTTTGTGCGGACAT	CGAGGAGATCGAGGCGGAGGAGGAGGA	2400
Db	2341	TTCCCCACATGAAGCCCTGTTTGTGCGGACAT	CGAGGAGATCGAGGCGGAGGAGGAGGAGGA	2400
Qy	2401	AGCGGAGCTGCGGAGGTGCGGGCGGCCCTCT	CTCTCAGGAGACTGGCAGGCGGCTCG	2460
Db	2401	AGCGGAGCTGCGGAGGTGCGGGCGGCCCTCT	CTCTCAGGAGACTGGCAGGCGGCTCG	2460
Qy	2461	AGGATGGGAGCCTCAGCAGAAAGCGGGCCCA	CACAGAGGAGCCACAGGCCAAAGAAGTCA	2520
Db	2461	AGGATGGGAGCCTCAGCAGAAAGCGGGCCCA	CACAGAGGAGCCACAGGCCAAAGAAGTCA	2520
Qy	2521	GAGCCAGTGAAGATCTGGGAGACCCCTGAACT	CAGAGAGGCTGTGTGTCTTCTGCCCCACG	2580
Db	2521	GAGCCAGTGAAGATCTGGGAGACCCCTGAACT	CAGAGAGGCTGTGTGTCTTCTGCCCCACG	2580
Qy	2581	CAGCCACCGTATCTGCCCTCTCTGCTGTGTA	GAGCTGTGAGAGCTGCGCAGGAGG	2640
Db	2581	CAGCCACCGTATCTGCCCTCTCTGCTGTGTA	GAGCTGTGAGAGCTGCGCAGGAGG	2640
Qy	2641	CAGCTCAGGATAGGTGGTATGAGCTGTGCCAG	GCTTGGGCTCCACATAAGCACTAGT	2700
Db	2641	CAGCTCAGGATAGGTGGTATGAGCTGTGCCAG	GCTTGGGCTCCACATAAGCACTAGT	2700
Qy	2701	CTATAGATGCTCTTAGGACTCGTCTGGCACAG	CAGCGCGGCGCAGAGGCTGCCACAG	2760
Db	2701	CTATAGATGCTCTTAGGACTCGTCTGGCACAG	CAGCGCGGCGCAGAGGCTGCCACAG	2760
Qy	2761	GAAGCAAGCAGATGAACCTAATTTCAATTTCA	AGGCGAGTTTTTAAAGAAAGTCTTGGAAA	2820
Db	2761	GAAGCAAGCAGATGAACCTAATTTCAATTTCA	AGGCGAGTTTTTAAAGAAAGTCTTGGAAA	2820
Qy	2821	ACGGCGCACCTTTCTTAATTCAGCAAAAGTGA	TTCCTGCACACACAGAGACAGCAGA	2880
Db	2821	ACGGCGCACCTTTCTTAATTCAGCAAAAGTGA	TTCCTGCACACACAGAGACAGCAGA	2880
Qy	2881	GTAACAGGATCAGTGGGCTCTAAGTGTCCGAG	ACTTAAACGAAAAAGTATTTTCAGCTGCAA	2940
Db	2881	GTAACAGGATCAGTGGGCTCTAAGTGTCCGAG	ACTTAAACGAAAAAGTATTTTCAGCTGCAA	2940
Qy	2941	TAAAGATTGAGTTTGCAA	2958	
Db	2941	TAAAGATTGAGTTTGCAA	2958	

GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
; US-09-988-626-223

Query Match 97.2%; Score 2874.4; DB 9; Length 2908;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY	51	ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGTCCGGCGCGGACGACCACTATGTCGCGAGGACGC 110
DB	1	ATGTGGCGCTTTGCTCGTGTGCTGCGTCCGGTCCGGCGCGGACGACCACTATGTCGCGAGGACGC 60
QY	111	ACCATATCGCAGCAGCCCGCCCGCGAGCGCGCGGACGCGGACCCGCTGGGACCTG 170
DB	61	ACCATATCGCAGCAGCCCGCCCGCGAGCGCGCGGACGCGGACCCGCTGGGACCTG 120
QY	171	CGCACCGGAGAGAACGCGGACCGTCCGGGTGCTCCGGCGGCGCCAAACACCGGTACCTG 230
DB	121	CGCACCGGAGAGAACGCGGACCGTCCGGGTGCTCCGGCGGCGCCAAACACCGGTACCTG 180
QY	231	CAGGTGGTGCAGCGGTAGCGGGACTCGGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 290
DB	181	CAGGTGGTGCAGCGGTAGCGGGACTCGGGCGCGCGCTCTACGCTCTTCTCCGAGTTC 240
QY	291	AACCGGTATCTTCAACTGTGTGAGAGCGGTTTCAGAGACTCATGTCAGGACACAAGTTA 350
DB	241	AACCGGTATCTTCAACTGTGTGAGAGCGGTTTCAGAGACTCATGTCAGGACACAAGTTA 300
QY	351	AAGTTGTCGCTGGACAACTATTCCTGACACGAATACGACTGGTCTAATGTTGGGGC 410
DB	301	AAGTTGTCGCTGGACAACTATTCCTGACACGAATACGACTGGTCTAATGTTGGGGC 360
QY	411	TTAAGTGGAAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
DB	361	TTAAGTGGAAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
QY	471	CTTCCAACTGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGGAAGA 530
DB	421	CTTCCAACTGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGGAAGA 480
QY	531	ATAGAACTGGCTGTGGGGCCCACTCTGCCAGATACGAGATGAACACCATGACAGTT 590
DB	481	ATAGAACTGGCTGTGGGGCCCACTCTGCCAGATACGAGATGAACACCATGACAGTT 540
QY	591	TACCAATCCCAATACACAGTGAACAGAGAGGAGGAAAGCAACCACTGGCAGAGTCCA 650
DB	541	TACCAATCCCAATACACAGTGAACAGAGAGGAGGAAAGCAACCACTGGCAGAGTCCA 600
QY	651	GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTTCGAATGAAAT 710

DB	601	GAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCAGTCAAAATGAAAAT 660
QY	711	GAGCCACACCTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 770
DB	661	GAGCCACACCTTCCACATGGTGTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
QY	771	GTAGCTTTCACTGTGTAAGCTTCACTTAAAGAGAGAAATCTTGTGGTGTCAAAAGCAAG 830
DB	721	GTAGCTTTCACTGTGTAAGCTTCACTTAAAGAGAGAAATCTTGTGGTGTCAAAAGCAAG 780
QY	831	GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 890
DB	781	GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 840
QY	891	GGCAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGGTGAAGAGCTGTGTACTCTCTCA 950
DB	841	GGCAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGGTGAAGAGCTGTGTACTCTCTCA 900
QY	951	GATCTGTGTGCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCCATC 1010
DB	901	GATCTGTGTGCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCCATC 960
QY	1011	TGTGAAATGCCACCTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCGCTGGCTTGGTG 1070
DB	961	TGTGAAATGCCACCTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCGCTGGCTTGGTG 1020
QY	1071	GTTTCACATGGGCGCCAGCATCTGTGCTGTGGACAGCATGACAGCAGTGGATGGAGAGG 1130
DB	1021	GTTTCACATGGGCGCCAGCATCTGTGCTGTGGACAGCATGACAGCAGTGGATGGAGAGG 1080
QY	1131	TTTGGGCTGACACCCAGCAGCTTGTGCTCTGAATGAGAACTGTGCTCAGTTTCACAACTTT 1190
DB	1081	TTTGGGCTGACACCCAGCAGCTTGTGCTCTGAATGAGAACTGTGCTCAGTTTCACAACTTT 1140
QY	1191	CGCAGCGCAAGATTTCAAAACCCAGCTCAACCTCATCCACCCGGAATCTTCCCTGCTGTC 1250
DB	1141	CGCAGCGCAAGATTTCAAAACCCAGCTCAACCTCATCCACCCGGAATCTTCCCTGCTGTC 1200
QY	1251	ACCAGTTTCGCTGTGAAGAGAGGGGCGCCACCTCAGTGTGCCATGTTTCAGGGTGAA 1310
DB	1201	ACCAGTTTCGCTGTGAAGAGAGGGGCGCCACCTCAGTGTGCCATGTTTCAGGGTGAA 1260
QY	1311	TGCTCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTTACT 1370
DB	1261	TGCTCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTTACT 1320
QY	1371	TGCAATCTGAGGAATTCATAGTTGAGGGGTGAGCTTCCCACTTCAGCAGAGCGTG 1430
DB	1321	TGCAATCTGAGGAATTCATAGTTGAGGGGTGAGCTTCCCACTTCAGCAGAGCGTG 1380
QY	1431	CAGGAGTACAGGAGGTGCGCAGGACGCGCCAGCCAGCAGAGAGAAAGTCCAGTAC 1490
DB	1381	CAGGAGTACAGGAGGTGCGCAGGACGCGCCAGCCAGCAGAGAGAAAGTCCAGTAC 1440
QY	1491	CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1550
DB	1441	CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
QY	1551	GCACACTTGTCAACATTAAGCCCGACAGCTCTGTGCTACTGGAATGTTGGTGGGACACA 1610
DB	1501	GCACACTTGTCAACATTAAGCCCGACAGCTCTGTGCTACTGGAATGTTGGTGGGACAG 1560
QY	1611	TTTGGGAGCTGTGCGCTCATTACGGACACAGGTGGACAGGCTCCTGGGACCCCTGGCT 1670
DB	1561	TTTGGGAGCTGTGCGCTCATTACGGACACAGGTGGACAGGCTCCTGGGACCCCTGGCT 1620
QY	1671	GCTGTGTTGTCTCCACCTGCAGCAGATCAACACCGGCTTCCCAAGTATCTTGTGTG 1730
DB	1621	GCTGTGTTGTCTCCACCTGCAGCAGATCAACACCGGCTTCTAAATATCTTGTGTG 1680
QY	1731	CAGAGAGAACCGCTTGGCATCTTTGGGAAAGCGCTTCACTCCCTTCTGCTGGTGTGCC 1790

Db 1681 CAGAGAGAACGAGCCTTGCCATCTTTTGGAAAGCCCTTTACCCCTTTTGCTGGTGTGCC 1740
Qy 1791 CCCAACCCAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCCAGTGCAGAGGTCCTGCAC 1850
Db 1741 CCCAACCCAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCCAGTGCAGAGGTCCTGCAC 1800
Qy 1851 CACATCAGTATGATTCTCGCCAAATGCCCTTACAGGAAGGGCTGAGATCTCCAGTCTCTCA 1910
Db 1801 CACATCAGTATGATTCTCGCCAAATGCCCTTACAGGAAGGGCTGAGATCTCCAGTCTCTCA 1860
Qy 1911 GTGAAAGATTGATCAGTTTCGCTGTGGCAACATGTGATTTTGGAAAGATTTTCAGACCTGT 1970
Db 1861 GTGAAAGATTGATCAGTTTCGCTGTGGCAACATGTGATTTTGGAAAGATTTTCAGACCTGT 1920
Qy 1971 CTGGTGGCAGCTCAAGCATCGCTTTGGCTGTGGCTGTGGTGCACACCTCTGGCTGGAAA 2030
Db 1921 CTGGTGGCAGCTCAAGCATCGCTTTGGCTGTGGCTGTGGTGCACACCTCTGGCTGGAAA 1980
Qy 2031 GTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
Db 1981 GTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
Qy 2091 ACCCTCCTGATACATCAAGCCACCCCTGGAAAGATGGTTTGGAAAGAGACAGTGGAAAAG 2150
Db 2041 ACCCTCCTGATACATCAAGCCACCCCTGGAAAGATGGTTTGGAAAGAGACAGTGGAAAAG 2100
Qy 2151 ACACACAGCACAACGTCCCAAGCCATCAGCTGGGGATGCGGATGAAACGGGAGTTTCATT 2210
Db 2101 ACACACAGCACAACGTCCCAAGCCATCAGCTGGGGATGCGGATGAAACGGGAGTTTCATT 2160
Qy 2211 ATGCTGTAACCACTTCAGCCAGGCTATGCCAAGGTCCCTCTTCAGCCGCCAATTCAGC 2270
Db 2161 ATGCTGTAACCACTTCAGCCAGGCTATGCCAAGGTCCCTCTTCAGCCGCCAATTCAGC 2220
Qy 2271 GAGAAAGTGGGAGTTGCCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTGCCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTTGCACAATG 2280
Qy 2331 CCCAAGCTGATTTCCGCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2390
Db 2281 CCCAAGCTGATTTCCGCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2340
Qy 2391 CGCAGGGAGAACGGGAGCTGCGGAGGTGCGGGCGGCCCTCTCTGCAGGGAGCTTGCA 2450
Db 2341 CGCAGGGAGAACGGGAGCTGCGGAGGTGCGGGCGGCCCTCTCTGCAGGGAGCTTGCA 2400
Qy 2451 GCGGCTGGAGGATGGGAGCTTCAGCAGAAAGCGGGGCCACACAGAGAGGCCACAGGCC 2510
Db 2401 GCGGCTGGAGGATGGGAGCTTCAGCAGAAACGGGGCCACACAGAGAGGCCACAGGCC 2460
Qy 2511 AAGAAAGTCAGAGCCAGTGAAGATCTGGGAGACCCCTGAACTCAGAAAGCTGTGTCTTT 2570
Db 2461 AAGAAAGTCAGAGCCAGTGAAGATCTGGGAGACCCCTGAAATTCAGAAAGCTGTGTCTTT 2520
Qy 2571 CTGCCCCACGACGACCCGATCTGCCCTCTCTGCTGTAGAAGCTGAAGAGCACGGTC 2630
Db 2521 CTGCCCCACGACGACCCGATCTGCCCTCTCTGCTGTAGAAGCTGAAGAGCACGGTC 2580
Qy 2631 CCCAGGAGGACGCTCAGGATAGGTGGTATGAGAGCTGTGCCAGGCTTTGGGCTCCACAT 2690
Db 2581 CCCAGGAGGACGCTCAGGATAGGTGGTATGAGAGCTGTGCCAAGGCTTTGGGCTCCACAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCCCTTTAGGACTGGTGGGACAGCCGGGGCCAGGAGG 2750
Db 2641 AAGCACTAGTCTATAGATGCCCTTTAGGACTGGTGGGACAGCCGGGGCCAGGAGG 2700
Qy 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAAGGCAAGTTTAAAGAGTC 2810
Db 2701 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAAGGCAAGTTTAAAGAGGC 2760
Qy 2811 TTGAAAACAGACGGGGACCTTTCTCTTAATCCAGCAAGGATTCCTGTCACACCCAGA 2870
Db 2761 TTGAAAACAGACGGGACGACCTTTCTCTTAATCCAGCAAGGATTCCTGTCACACCCAGA 2820

Qy 2871 GACAAGCAGAGTACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATT 2930
Db 2821 GACAAGCAGAGTACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATT 2880
Qy 2931 TCAGCTCAATAAAGATTGAGTTTGCAA 2958
Db 2881 TCAGCTCAATAAAGATTGAGTTTGCAA 2908

RESULT 4
US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223

Query Match 97.2%; Score 2874.4; DB 9; Length 2908;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 51 ATGTGGGCGCTTTGCTCGCTCGCTCCGGCCGGACGCCACCATGTGCGAGGACGC 110
Db 1 ATGTGGGCGCTTTGCTCGCTCGCTCCGGCCGGACGCCACCATGTGCGAGGGACGC 60

Qy 111 ACCATATCGCAGGACCCGCCCGCGAGCGCGCCGCAAGGACCCGCTGCGGCACCTG 170
Db 61 ACCATATCGCAGGACCCGCCCGCGAGCGCGCCGCAAGGACCCGCTGCGGCACCTG 120

Qy 171 GGCACGCGAGAGAACCGCGACCGTGGGGTGTCCGGCGGCCCAACACCGTGTACTCTG 230
Db 121 GGCACGCGAGAGAACCGCGACCGTGGGGTGTCCGGCGGCCCAACACCGTGTACTCTG 180

Qy 231 CAGTGTGGCAGCGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTTCTTCCGAGTTC 290
Db 181 CAGTGTGGCAGCGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTTCTTCCGAGTTC 240

Qy 291 AACCGGTATCTTCAACTGTGGAGAAGCGTTTCAGAGACTCATGCAAGGACCAAGTTA 350
Db 241 AACCGGTATCTTCAACTGTGGAGAAGCGTTTCAGAGACTCATGCAAGGACCAAGTTA 300

Qy 351 AAGGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGC 410
Db 301 AAGGTGCTCGCTGGACAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGC 360

Qy 411 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 470
Db 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 420

QY 471 CTTCCAACTGGAATAATACTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGA 530
Db 421 CTTCCAACTGGAATAATACTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGA 480
QY 531 ATAGAACTGGCTGTGGGCCCCCACTCTGCCAGATACGAGGATGAACACCATGACAGTT 590
Db 481 ATAGAACTGGCTGTGGGCCCCCACTCTGCCAGATACGAGGATGAACACCATGACAGTT 540
QY 591 TACCAGATCCCAATACACAGTGAACAGAGAGGGGAAAGCAACCAATGGCAGAGTCCA 650
Db 541 TACCAGATCCCAATACACAGTGAACAGAGAGGGGAAAGCAACCAATGGCAGAGTCCA 600
QY 651 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGATCCAGAGTGAATAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGATCCAGAGTGAATAAAT 660
QY 711 GAGCCACACTTCCACATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 770
Db 661 GAGCCACACTTCCACATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
QY 771 GTAGCTTTTCATCTGTAACTTCACTTAAAGAGAGGAAACTTCTGGTGTCTCAAGCAAG 830
Db 721 GTAGCTTTTCATCTGTAACTTCACTTAAAGAGAGGAAACTTCTGGTGTCTCAAGCAAG 780
QY 831 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCTCTCTCTCAAGGAC 890
Db 781 GAGATGGGCTCCAGTGGGACAGCTGCCATCGCTCCCATCTCTCTCTCAAGGAC 840
QY 891 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 950
Db 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTGGCTGAAGAGCTGTACTCTCTCA 900
QY 951 GATCTGGTGTCTTTTGTGGTGTAGATGTCCAGATGAAGCTTCATTCAACCCATC 1010
Db 901 GATCTGGTGTCTTTTGTGGTGTAGATGTCCAGATGAAGCTTCATTCAACCCATC 960
QY 1011 TGTGAATGCCCCCTTTGAGAGTACCAAGGAAAGGAGATGCCCCCTGGCTGGTG 1070
Db 961 TGTGAATGCCCCCTTTGAGAGTACCAAGGAAAGGAGATGCCCCCTGGCTGGTG 1020
QY 1071 GTTCACTGGCCCCCAGATCTGTCTGTGGAGAGAGGTACAGAGCTGGATGGAGAG 1130
Db 1021 GTTCACTGGCCCCCAGATCTGTCTGTGGAGAGAGGTACAGAGCTGGATGGAGAG 1080
QY 1131 TTTGGGCTTGACACCCAGACTTGTCTTGAATGAGAACTGTGCCCTCAGTTCAACACTT 1190
Db 1081 TTTGGGCTTGACACCCAGACTTGTCTTGAATGAGAACTGTGCCCTCAGTTCAACACTT 1140
QY 1191 CGCAGCCAAAGATTCAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1250
Db 1141 CGCAGCCAAAGATTCAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
QY 1251 ACCAGTTTCGCTGTGAAGAGAGGGCCCCCCTCAGTGTGCCATGGTTGAGGTGAA 1310
Db 1201 ACCAGTTTCGCTGTGAAGAGAGGGCCCCCCTCAGTGTGCCATGGTTGAGGTGAA 1260
QY 1311 TGCCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1370
Db 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTATTACT 1320
QY 1371 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGAGAGCGTG 1430
Db 1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCACTTCCAGAGAGCGTG 1380
QY 1431 CAGGAGTACGAGAGAGTGGCAGGACGCGCCAGCGCCAGAGAGAGAGAGAGTCAAGTAC 1490
Db 1381 CAGGAGTACGAGAGAGTGGCAGGACGCGCCAGCGCCAGAGAGAGAGAGAGTCAAGTAC 1440
QY 1491 CCAGAAATCATCTTCTTGGAAAGAGGTCTGCCATCCCGATGAGATTCGAATGTCACT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAAGAGGTCTGCCATCCCGATGAGATTCGAATGTCACT 1500

QY 1551 GCCACACTTGTCAACATTAAGCCCCGACACAGCTCTCTGTCTACTGGACTGTGGTGAAGGACA 1610
Db 1501 GCCACACTTGTCAACATTAAGCCCCGACACAGCTCTCTGTCTACTGGACTGTGGTGAAGGACG 1560
QY 1611 TTTGGGAGCTGTGCGCTCATTTACGGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1670
Db 1561 TTTGGGAGCTGTGCGCTCATTTACGGAGACACAGGTGGACAGGCTCTGGGACACCTGGCT 1620
QY 1671 GCTGTGTTGTGTCCACTCTGCACGCAGATCAACACCGGCTTCCCAAGTATCTTGTGTG 1730
Db 1621 GCTGTGTTGTGTCCACTCTGCACGCAGATCAACACCGGCTTCCCAAGTATCTTGTGTG 1680
QY 1731 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCGCTTACCCCTTGTGTGTGTGCT 1790
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCGCTTACCCCTTGTGTGTGTGCT 1740
QY 1791 CCCAACACAGCTCAAAAGCTGGCTCCAGCAGTACACAAACAGTGCAGAGGTCTCTGCAC 1850
Db 1741 CCCAACACAGCTCAAAAGCTGGCTCCAGCAGTACACAAACAGTGCAGAGGTCTCTGCAC 1800
QY 1851 CACATCAGTATGATCTCTGCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTCTCA 1910
Db 1801 CACATCAGTATGATCTCTGCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTCTCA 1860
QY 1911 GTGGAAGATGATCAGTTCGCTGTGGAAACATGTGATTTGGAAAGATTTTCAGACCTGT 1970
Db 1861 GTGGAAGATGATCAGTTCGCTGTGGAAACATGTGATTTGGAAAGATTTTCAGACCTGT 1920
QY 1971 CTGGTGGCGCATGCAAGCATCGTTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 2030
Db 1921 CTGGTGGCGCATGCAAGCATCGTTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGAA 1980
QY 2031 GTGGTCTATTTCGGGGACACCATGCTCCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2090
Db 1981 GTGGTCTATTTCGGGGACACCATGCTCCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
QY 2091 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGAAAG 2150
Db 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGAAAG 2100
QY 2151 ACACACAGCACAACCTGCCAAGCCATCAGCGTGGGATGCCGATGAACGCGAGTTCATT 2210
Db 2101 ACACACAGCACAACCTGCCAAGCCATCAGCGTGGGATGCCGATGAACGCGAGTTCATT 2160
QY 2211 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTTCCTCTTCAGCCCCCACTTCAGC 2270
Db 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTTCCTCTTCAGCCCCCACTTCAGC 2220
QY 2271 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGATCTTTCACCAATG 2330
Db 2221 GAGAAAGTGGGAGTTCCTTTGACCAATGAAGTCTGCTTTGGAGATCTTTCACCAATG 2280
QY 2331 CCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTGTGGGAGATCGAGGATGAGGAG 2390
Db 2281 CCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTGTGGGAGATCGAGGATGAGGAG 2340
QY 2391 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCGCTCTCTTCAGGAGAGTGGCA 2450
Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCGCTCTCTTCAGGAGAGTGGCA 2400
QY 2451 GCGGCTCTGGAGATGGGAGCTTCAGAGAAAGCGGCGCCACACAGAGAGAGCCAGGCT 2510
Db 2401 GCGGCTCTGGAGATGGGAGCTTCAGAGAAAGCGGCGCCACACAGAGAGAGCCAGGCT 2460
QY 2511 AAGAGGTCAGAGCCCACTGAAGATCTGGGAGACCTGAACTCAGAGGCTGTGTCTT 2570
Db 2461 AAGAGGTCAGAGCCCACTGAAGATCTGGGAGACCTGAACTCAGAGGCTGTGTCTT 2520
QY 2571 CTGCCCCCAGCAGCCCTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2630
Db 2521 CTGCCCCCAGCAGCCCTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580
QY 2631 CCCCAGGAGGAGCTCAGGATAGGTGGTATGAGGCTGTGCCGAGGCTTGGGCTCCACAT 2690

Db 2581 CCCAGGAGGAGCTCAGGATAGGTGGTATGAGCTGTGCCAAGGCTTGGGCTCCCAAT 2640
Qy 2691 AAGCACTAGTCTATAGATGCTCTTAGGACTGGTGCCTGGGCAAGCCGGGCCAGGAGG 2750
Db 2641 AAGCACTAGTCTATAGATGCTCTTAGGACTGGTGCCTGGGCAAGCCGGGCCAGGAGG 2700
Qy 2751 CTGCCACAGGAGCAAGCAGATGAACCTTAATTTTCAATTTCAAGGCGAGTTTAAAGAGTC 2810
Db 2701 CTGCCACAGGAGCAAGCAGATGAACCTTAATTTTCAATTTCAAGGCGAGTTTAAAGAGGC 2760
Qy 2811 TTGGAAACAGACCGGGGACCTTCTCTTAATCCAGCAAGTGAATTCCTCTGCACACCAGA 2870
Db 2761 TTGGAAACAGACCGGGGACCTTCTCTTAATCCAGCAAGTGAATTCCTCTGCACACCAGA 2820
Qy 2871 GACAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAAACGAAATAGTATT 2930
Db 2821 GACAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAAACGAAATAGTATT 2880
Qy 2931 TCAGCTGCAGTAAGATTCAGTTTGCAA 2958
Db 2881 TCAGCTGCAGTAAGATTCAGTTTGCAA 2908

RESULT 5

US-09-988-626-225
; Sequence 225, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)

US-09-988-626-225

Query Match 95.3%; Score 2819.6; DB 9; Length 2892;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 2863; Conservative 0; Mismatches 29; Indels 16; Gaps 1;
Qy 51 ATGTGGGGCTTTGCTCGTGTCTCGGTCCGGCCGGAGCGGAGCCATGTTCGAGGAGCGC 110
Db 1 ATGTGGGGCTTTGCTCGTGTCTCGGTCCGGCCGGAGCGGAGCCATGTTCGAGGAGCGC 60
Qy 111 ACCATATCGCAGGACCCCGCCCGGAGCGGCGGAGGACCCGCTCGGGACCTG 170
Db 61 ACCATATCGCAGGACCCCGCCCGGAGCGGCGGAGGACCCGCTCGGGACCTG 120
Qy 171 CGCAGCGAGAGAGCGGAGCGGAGCGGCGGAGGACCCGCTCGGGAGCGGACCGTGTACCTG 230
Db 121 CGCAGCGAGAGAGCGGAGCGGAGCGGCGGAGGACCCGCTCGGGAGCGGACCGGCTGTACCTG 180
Qy 231 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCGGCGCTCTACGTCTTCTCCGAGTTC 290

Db 181 CAGTGTGGCAGCGGTAGCCGGACTCGGGCGCCGGCTCTACGTCTTCTCCAGTTC 240
Qy 291 AACCGGTATCTTTCAACTGTGGAGAGGGCTTCAGAGACTCATGACAGGACCAAGTTA 350
Db 241 AACCGGTATCTTTCAACTGTGGAGAGGGCTTCAGAGACTCATGACAGGACCAAGTTA 300
Qy 351 AAGTTTCTCGCTGGGCAACATATTCCTCACAGGAATGCACTGGTCTAATGTGGGGC 410
Db 301 AAGTTTCTCGCTGGGCAACATATTCCTCACAGGAATGCACTGGTCTAATGTGGGGC 360
Qy 411 TTAAGTGAATGATTCTTACTTTAAAGAAAACCGGGCTTCCAAAGTGTGTACTTTCTTGA 470
Db 361 TTAAGTGAATGATTCTTACTTTAAAGAAAACCGGGCTTCCAAAGTGTGTACTTTCTTGA 420
Qy 471 CCTCCAACTTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGTCTCAATGAAAGGA 530
Db 421 CCTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGTCTCAATGAAAGGA 480
Qy 531 ATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCAACCAACCATGGCAGAGTCCA 650
Db 541 TACCAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCAACCAACCATGGCAGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAATGAAAT 660
Qy 711 GAGCCACACTTCCACATGTGTAGCCAGAGAGGGGTGAGGAGTCTTCCCTGGTC 770
Db 661 GAGCCACACTTCCACATGTGTAGCCAGAGAGGGGTGAGGAGTCTTCCCTGGTC 720
Qy 771 GTAGCTTTTCTGTAAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCTCAAGCAAG 830
Db 721 GTAGCTTTTCTGTAAAGCTTCACTTAAAGAGAGAACTTCTTGGTGTCTCAAGCAAG 780
Qy 831 GAGATGGGCTCCAGTTGGGACAGTGCATCGCTCCCATCATTTGTGTGTCAAGAC 890
Db 781 GAGATGGGCTCCAGTTGGGACAGTGCATCGCTCCCATCATTTGTGTGTCAAGAC 840
Qy 891 GGGAAAGCATCACTCATGAAGNAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 950
Db 841 GGGAAAGCATCACTCATGAAGNAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 951 GATCCTGTGTCTGCTTTTGTGGTGAATGTCCAGATGAAAGCTTCAATCAACCCATC 1010
Db 901 GATCCTGTGTCTGCTTTTGTGGTGAATGTCCAGATGAAAGCTTCAATCAACCCATC 960
Qy 1011 TGTGAGAAATGCCACTTTCAGAGTCAAGGAAAGCAGATGCCCGGTGGCTTGGT 1070
Db 961 TGTGAGAAATGCCACTTTCAGAGTCAAGGAAAGCAGATGCCCGGTGGCTTGGT 1020
Qy 1071 GTTCACATGGCCCGCAGCATCTGTGCTTGTGGACAGAGTACCAGAGTGGATGGAGG 1130
Db 1021 GTTCACATGGCCCGCAGAAATCTGTGCTTGTGGACAGAGTACCAGAGTGGATGGAGG 1080
Qy 1131 TTTGGGCTTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAACACTT 1190
Db 1081 TTTGGGCTTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAACACTT 1140
Qy 1191 CGCAGCACAAGATTCAAAACCCAGCTCAACCTCATCCCGGAGACATCTTCCCGCTGCT 1250
Db 1141 CGCAGCACAAGATTCAAAACCCAGCTCAACCTCATCCCGGAGACATCTTCCCGCTGCT 1200
Qy 1251 ACCAGTTTCCGCTGAAGAGAGGGGCCCCACCTCAGTGTGCCATGGTTTCAAGGTGAA 1310
Db 1201 ACCAGTTTCCGCTGAAGAGAGGGGCCCCACCTCAGTGTGCCATGGTTTCAAGGTGAA 1260
Qy 1311 TGCCTCTCAAGTACCAAGCTCCGTCCCGAGGAGGTGGCAGAGGATGCCATTATTACT 1370

Db 1261 TGCCTCCTCAAGTACCAGCTCCGTCCAGAGGGAAATGCGAGAGGATGCCATTATCACT 1320
QY 1371 TGAATCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAATTCACAGAGCGTG 1430
Db 1321 TGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAATTCACAGAGTG 1380
QY 1431 CAGGAGTACAGGAGGAGTGGCAGGACGCGCCAGCCGACGAGAGAGAAAGAGTCAGTAC 1490
Db 1381 CAGGAGTACAGGAGGAGTGGCAGGACGCTCCAGCCGACGAGAGAGAAAGAGTCAGTAC 1440
QY 1491 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
QY 1551 GCCACACTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGCACTGTGGTGAAGGACACA 1610
Db 1501 GCCACACTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGCACTGTGGTGAAGGACAG 1560
QY 1611 TTTGGGACGCTGCGGTCATTACGAGACACAGGTGGAGAGGTCCTGGGACCCCTGGCT 1670
Db 1561 TTTGGGACGCTGCGGTCATTACGAGACACAGGTGGAGAGGTCCTGGGACCCCTGGCT 1620
QY 1671 GCTGTCTTGTGTCACCTGCAAGCAGATCACACACGGCTTGCCAAAGTATCTTGTCTG 1730
Db 1621 GCTGTCTTGTGTCACCTGCAAGCAGATCACACACGGCTTGCCAAAGTATCTTGTCTG 1680
QY 1731 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCCTTTCACCTTTGCTGTGGTGTGCC 1790
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTGGGAAAGCCCTTTCACCTTTGCTGTGGTGTGCC 1740
QY 1791 CCCAACGCTTAAAGCCTGGCTCAGAGTACCAACACAGTGCAGAGGTCCTGTGCAC 1850
Db 1741 CCCAGCAGCTCAAAAGCCTGGCTCAGAGTACCAACACAGTGCAGAGGTCCTGTGCAC 1800
QY 1851 CACATCAGTATGATCTCGCAATTCCTTTCAGGAGGCGCTGAGATCTCCAGTCTCTGCA 1910
Db 1801 CACATCAGTATGATCTCGCAATTCCTTTCAGGAGGCGCTGAGATCTCCAGTCTCTGCA 1860
QY 1911 GTGGAAGATTCATAGTTTCGCTGTTCGCAACATGTGATTTGGAAGATTTTCAGACCTGT 1970
Db 1861 GTGGAAGATTCATAGTTTCGCTGTTCGCAACATGTGATTTGGAAGATTTTCAGACCTGT 1920
QY 1971 CTGGTCGGGACTGCAAGCATGGTTTGGCTGTGCTGCTGGTGCACACCTCTGGCTGGA 2030
Db 1921 CTGGTCGGGACTGCAAGCATGGTTTGGCTGTGCTGCTGGTGCACACCTCTGGCTGGA 1980
QY 2031 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGCTCGGATGGGAAAGATGCC 2090
Db 1981 GTGGTCTATTCCGGGACACCATGCTCGAGGCTCTGCTCGGATGGGAAAGATGCC 2040
QY 2091 ACCCTCCTGATACATGAAGCCACCTCGGAAGATGGTTTGGAGAGGAAGCAGTGGAAAG 2150
Db 2041 ACCCTCCTGATACATGAAGCCACCTCGGAAGATGGTTTGGAGAGGAAGCAGTGGAAAG 2100
QY 2151 ACACAGCACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGGATTCATT 2210
Db 2101 ACACAGCACAACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACGCGGATTCATT 2160
QY 2211 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAGC 2270
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAGC 2220
QY 2271 GAGAAAGTGGGAGTTGCTTTCACCACTGAAGTCTGTCTTGGAGACTTTTCCAAACATG 2330
Db 2221 GAGAAAGTGGGAGTTGCTTTCACCACTGAAGTCTGTCTTGGAGACTTTTCCAAACATG 2280
QY 2331 CCCAAGCTGATTCCTCCACTGAAGCCCTGTTTGTGGCGACATCGAGGAGATGGAGGAG 2390
Db 2281 CCCAAGCTGATTCCTCCACTGAAGCCCTGTTTGTGGCGACATCGAGGAGATGGAGGAG 2340
QY 2391 CGCAGGAGAGCGGAGCTGCGGCAAGTTCGGGCGGCGCTCTCTGTCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAGCGGAGCTGCGGCAAGTTCGGGCGGCGCTCTCTGTCAGGAGCTGGCA 2400

QY 2451 GCGGCGCTGGAGGATGGGAGCGCTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGGAGGATGGGAGCGCTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
QY 2511 AAGAGGTCAGAGGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTGCTT 2570
Db 2461 AAGAGGTCAGAGGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTGCTT 2520
QY 2571 CTGCCCCACGACGACCGCTATCTGCCCTCTCTTCTGCTGTAGAACTGAAGACGCGTC 2630
Db 2521 CTGCCCCACGACGACCGCTATCTGCCCTCTCTTCTGCTGTAGAACTGAAGACGCGTC 2580
QY 2631 CCCAGAGGAGGAGCTCAGGATAGGTGGTATGAGGCTGTGCCAGGCTTGGGCTCCACAT 2690
Db 2581 CCCAGAGGAGGAGCTCAGGATAGGTGGTATGAGGCTGTGCCAGGCTTGGGCTCCACAT 2640
QY 2691 AAGCACTAGTCTATAGATGCCCTCTTAGGACTGGTCTGGCAGCAGCGCGGCGCAGGAGG 2750
Db 2641 AAGCACTAGTCTATTA-----GGTGCCTGGCAGCAGCGCGGCGCAGGAGG 2684
QY 2751 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCGATTTTAAAGAAAGTC 2810
Db 2685 CTGCCACACGGAAGCAAGCAGATGAATTAATTTCAATTTCAAGGCGATTTTAAAGAAAGTC 2744
QY 2811 TTGGAACACAGCGCGGACCTTCTCTAATCCAGCAAAAGTGATTCCTGTCACACACAGA 2870
Db 2745 TTGGAACACAGCGCGGACCTTCTCTAATCCAGCAAAAGTGATTCCTGTCACACACAGA 2804
QY 2871 GACAAGCAGATTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAAGCAAAATAGTATT 2930
Db 2805 GACAAGCAGATTAACAGATCAGTGGGTCTAAGTGTCCGAGACTTAAAGCAAAATAGTATT 2864
QY 2931 TCAGCTGCAATAAAGATTGAGTTTGCAA 2958
Db 2865 TCAGCTGCAATAAAGATTGAGTTTGCAA 2892

RESULT 6

US-09-988-687-225
; Sequence 225, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-225

Query Match 95.3%; Score 2819.6; DB 9; Length 2892;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 2863; Conservative 0; Mismatches 29; Indels 16; Gaps 1;

Qy 51 ATGTGGCGCTTTGCTCGCTGCTGGTCCGCGCGGACGACCATGTCCGAGGGACG 110
Db 1 ATGTGGCGCTTTGCTCGCTGCTGGTCCGCGCGGACGACCATGTCCGAGGGACG 60
Qy 111 ACCATATCGCAGCACCCGCGCGCGAGCGCGCGCAAGACCCGCTGCGGCACCTG 170
Db 61 ACCATATCGCAGCACCCGCGCGCGAGCGCGCGCAAGACCCGCTGCGGCACCTG 120
Qy 171 CGCACGAGAGAACGGCGGACCTGCGGGTGTCTCGCGCGGCCAAAACCGGTGTACTG 230
Db 121 CGCACGAGAGAACGGCGGACCTGCGGGTGTCTCGCGCGGCCAAAACCGGTGTACTG 180
Qy 231 CAGGTGTGGCAGCGGTAGCCGGACCTCGCGCGCGCGCTCTAGCTCTTCTCCAGTTC 290
Db 181 CAGGTGTGGCAGCGGTAGCCGGACCTCGCGCGCGCGCTCTAGCTCTTCTCCAGTTC 240
Qy 291 AACCGGTATCTCTCAACTGTGGAGGCGTTCCAGAGACTCATCGCAGGACACAAGTTA 350
Db 241 AACCGGTATCTCTCAACTGTGGAGAGGCGTTCCAGAGACTCATCGCAGGACACAAGTTA 300
Qy 351 AAGTTGTCTGCTGGACAAATATCTCTGACACGAATGCACTGTCTTAATGTTGGGGC 410
Db 301 AAGTTGTCTGCTGGACAAATATCTCTGACACGAATGCACTGTCTTAATGTTGGGGC 360
Qy 411 TTAAGTGGNAATGTTCTTACTTTAAAGNAACCGGGCTTCCAAAGTGTACTTTCTGGA 470
Db 361 TTAAGTGGNAATGTTCTTACTTTAAAGNAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Qy 471 CTTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGGTCCATTGAAAGGA 530
Db 421 CTTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGGTCCATTGAAAGGA 480
Qy 531 ATAGAACTGGTGTGGGGCCCACTCTGCCCCAGAAATACGAGGATGAAACCATGACGTT 590
Db 481 ATAGAACTGGTGTGGGGCCCACTCTGCCCCAGAAATACGAGGATGAAACCATGACGTT 540
Qy 591 TACCAATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCACTGCGAGTCCA 650
Db 541 TACCAATCCCATACACAGTGAACAGAGAGGGGAAAGCAACCACTGCGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCCAGCTCCGAGTCCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCCAGCTCCGAGTCCGAATGAAAT 660
Qy 711 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGTC 770
Db 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGTC 720
Qy 771 GTAGCTTTTCATCTGAAGCTTCACTTAAAGAGAGAACTTTTGGTGTCAAAAGCAAG 830
Db 721 GTAGCTTTTCATCTGAAGCTTCACTTAAAGAGAGAACTTTTGGTGTCAAAAGCAAG 780
Qy 831 GAGATGGGCTCCAGTTGGGACAGCTGCATCGCTCCCATCATTTGCTGTCTCAAGAC 890
Db 781 GAGATGGGCTCCAGTTGGGACAGCTGCATCGCTCCCATCATTTGCTGTCTCAAGAC 840
Qy 891 GGGAAAGCATCACTCATGAAGGAAGAGATTTTGGGTGAAGAGCTGTACTCTCTCA 950
Db 841 GGGAAAGCATCACTCATGAAGGAAGAGATTTTGGGTGAAGAGCTGTACTCTCTCA 900
Qy 951 GATCTGTGTGCTTTTGTGGTGTAGAACTCCAGATGAAGCTTCAATTCACCCCATC 1010
Db 901 GATCTGTGTGCTTTTGTGGTGTAGAACTCCAGATGAAGCTTCAATTCACCCCATC 960
Qy 1011 TGTGAGATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCGCTGGCTTGTG 1070
Db 961 TGTGAGATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCGCTGGCTTGTG 1020
Qy 1071 GTTTCATGGCCCCAGCATCTGTGTTGTGGAACAGAGGTACCAAGTGTGAAGGAGG 1130
Db 1021 GTTTCATGGCCCCAGCATCTGTGTTGTGGAACAGAGGTACCAAGTGTGAAGGAGG 1080

Qy 1131 TTTGGGCTGACACCCAGCACTTGGTCTGAAATGAGAACTGTGCTCAGTTCAACACTT 1190
Db 1081 TTTGGGCTGACACCCAGCACTTGGTCTGAAATGAGAACTGTGCTCAGTTCAACACTT 1140
Qy 1191 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1250
Db 1141 CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTC 1200
Qy 1251 ACAGTTTCCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCATGGTTTCAGGGTGA 1310
Db 1201 ACAGTTTCCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCATGGTTTCAGGGTGA 1260
Qy 1311 TGCCTCTCAAGTACCACTCCGTCAGAGGAGGAGTGGCAGAGGATGCCATTATTACT 1370
Db 1261 TGCCTCTCAAGTACCACTCCGTCAGAGGAGGATGGCAGAGGATGCCATTATTACT 1320
Qy 1371 TGCAAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGTGTG 1430
Db 1321 TGCAAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGTGTG 1380
Qy 1431 CAGGAGTACAGGAGGAGTGGCAGGACGGCCCCAGCCCCAGAGAGAGAAAGTCACTAC 1490
Db 1381 CAGGAGTACAGGAGGAGTGGCAGGACGGCCCCAGCCCCAGAGAGAGAAAGTCACTAC 1440
Qy 1491 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCCGAAATGTCTAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCCGAAATGTCTAGT 1500
Qy 1551 GCCACACTTGTCAACATAAAGCCCCGACACGCTCTCTGTACTTGGACTGTGGTGGGGCACA 1610
Db 1501 GCCACACTTGTCAACATAAAGCCCCGACACGCTCTCTGTACTTGGACTGTGGTGGGGCACA 1560
Qy 1611 TTTGGGAGCTGTGCGTCAATTACGAGACAGGTGACAGGGTCTTGGGACCCCTGGCT 1670
Db 1561 TTTGGGAGCTGTGCGTCAATTACGAGACAGGTGACAGGGTCTTGGGACCCCTGGCT 1620
Qy 1671 GCTGTGTTGTGTCACACTGCGACAGAGTACCACAGGGCTTGGCAAGTATCTTCTGCTG 1730
Db 1621 GCTGTGTTGTGTCACACTGCGACAGAGTACCACAGGGCTTGGCAAGTATCTTCTGCTG 1680
Qy 1731 CAGAGAAACGCGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGC 1790
Db 1681 CAGAGAAACGCGCTTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGGTGGC 1740
Qy 1791 CCCAACAGCTCAAGCTGGCTCCAGAGTACCAACAGTCCAGGAGGCTCTGCAC 1850
Db 1741 CCCAGCAGCTCAAGCTGGCTCCAGAGTACCAACAGTCCAGGAGGCTCTGCAC 1800
Qy 1851 CACATCAGTATGATTCTGTCGCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTTGCA 1910
Db 1801 CACATCAGTATGATTCTGTCGCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTTGCA 1860
Qy 1911 GTGGAAAGATTGATCAGTTCCGCTGTTGCGAAACATGATTTTGAAGAGTTTCAGACCTGT 1970
Db 1861 GTGGAAAGATTGATCAGTTCCGCTGTTGCGAAACATGATTTTGAAGAGTTTCAGACCTGT 1920
Qy 1971 CTGCTGGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACTCTGGCTGAAA 2030
Db 1921 CTGCTGGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACTCTGGCTGAAA 1980
Qy 2031 GTGCTCTATTTCGGGGGACACATGCCCTGCGAGGCTCTGCTGCCGATGGGAAAGATGCC 2090
Db 1981 GTGCTCTATTTCGGGGGACACATGCCCTGCGAGGCTCTGCTGCCGATGGGAAAGATGCC 2040
Qy 2091 ACCCTCTGATACATGAAGCCACCTCGAAAGATGTTTGAAGAGGAGCAGTGGGAAAG 2150
Db 2041 ACCCTCTGATACATGAAGCCACCTCGAAAGATGTTTGAAGAGGAGCAGTGGGAAAG 2100
Qy 2151 ACACAGCACAAGCTTCCAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTTATT 2210
Db 2101 ACACAGCACAAGCTTCCAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTTATT 2160
Qy 2211 ATGCTGAACCACTTCAGCCAGCGCTATGTCGCAAGGTCCCCCTCTTTCAGCCCCCACTTCAGC 2270


```
Db 2161 ATGCTGAACCACTTCAGCCAGGCTATGCCAAGGTCCTCTTCCAGCCCCAACTTCAAC 2220
Qy 2271 GAGAAAGTGGGAGTTCCTTTTACCACATGAAGGCTGCTTTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTCCTTTTACCACATGAAGGCTGCTTTTGGAGACTTTTCCAAACAATG 2280
Qy 2331 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGAGCATCAGAGAGATGAGAGG 2390
Db 2281 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGGGAGCATCAGAGAGATGAGAGG 2340
Qy 2391 CCAGGGAGAACGGGAGCTCGGAGGTGCGGAGGTGCGGCGGCCCTCTGTCACAGGAGCTGGCA 2450
Db 2341 CCAGGGAGAACGGGAGCTCGGAGGTGCGGAGGTGCGGCGGCCCTCTGTCGCGGAGCTGGCA 2400
Qy 2451 GCGGCGCTGGAGATGGGAGGCTTCAGCAGAAAGCGGGCCCAACACAGAGAGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGGAGATGGGAGGCTTCAGCAGAAACGGGCCCAACACAGAGAGAGCCACAGGCC 2460
Qy 2511 AAGAAGGTGAGAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTT 2570
Db 2461 AAGAAGGTGAGAGCCAGTGAAGATCTGGGAGACCTGAACTCAGAAAGCTGTGTCTT 2520
Qy 2571 CTGCCCCCAGCAGCCAGCTATCTGCCCTCTTCTGCTGAGAGCTGAAGCAGCGTC 2630
Db 2521 CTGCCCCCAGCAGCCAGCTATCTGCCCTCTTCTGCTGAGAGCTGAAGCAGCGTC 2580
Qy 2631 CCCCAGGAGGAGCTCAGATAGGTGATGAGAGCTGCCAGGCTTGGGCTCCACAT 2690
Db 2581 CCCCAGGAGGAGCTCAGATAGGTGATGAGAGCTGCCAGGCTTGGGCTCCACAT 2640
Qy 2691 AAGCACTAGCTATAGATGCTCTTAGGACTGGTGGCTGGCAAGCCGCGGCGCAGGAGG 2750
Db 2641 AAGCACTAGCTATATA-----GGTGCCTGGCAGCGCGGCGCAGGAGG 2684
Qy 2751 CTGCCACAGGAGCAAGCAGATGACTTAATTTCAATTTCAAGGCAGTTTTAAAGAGTC 2810
Db 2685 CTGCCACAGGAGCAAGCAGATGACTTAATTTCAATTTCAAGGCAGTTTTAAAGAGTC 2744
Qy 2811 TTGGAACACAGCGGGGACCTTCTCTTAATCCAGCAAGTGATTCCTGTCACACCAGA 2870
Db 2745 TTGGAACACAGCGGAGCAGCTTCTCTTAATCCAGCAAGTGATTCCTGTCACACCAGA 2804
Qy 2871 GACAGCAGAGTAACAGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATT 2930
Db 2805 GACAGCAGAGTAACAGATCAGTGGGTCTAAGTGTCCAGACTTAACGAAATAGTATT 2864
Qy 2931 TCAGCTGCAATAAAGATTGAGTTGCAA 2958
Db 2865 TCAGCTGCAATAAAGATTGAGTTGCAA 2892
```

RESULT 7

```
US-09-988-626-1
; Sequence 1, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; TITLE OF INVENTION: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
```

```
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-1
```

```
Query Match 83.9%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 51 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGCGCGGAGCACCACCATGTGCGAGGAGCGC 110
Db 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCGCGCGCGGAGCACCACCATGTGCGAGGAGCGC 60
Qy 111 ACCATATCCAGAGCAACCCGCGCGCGCGAGCGCGCGCCAAAGGACCCGCTGCGGACCTG 170
Db 61 ACCATATCCAGAGCAACCCGCGCGCGCGAGCGCGCGCCAAAGGACCCGCTGCGGACCTG 120
Qy 171 GCGAGCGAGAGAGCGCGGAGCCGTCGCGGTCGCTCGGCGCGCGCCAAACACCGCTGACCTG 230
Db 121 GCGAGCGAGAGAGCGCGGAGCCGTCGCGGTCGCTCGGCGCGCGCCAAACACCGCTGACCTG 180
Qy 231 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGCGCGCTTACGCTCTTCTCCGAGTTC 290
Db 181 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGCGCGCTTACGCTCTTCTCCGAGTTC 240
Qy 291 AACCGGTATCTTCAACTGTGGAGAGCGGCTTACAGAGACTCATCGAGAGACACAAGTTA 350
Db 241 AACCGGTATCTTCAACTGTGGAGAGCGGCTTACAGAGACTCATCGAGAGACACAAGTTA 300
Qy 351 AAGTTGCTCGCTGGACAAACATATCTCTGACAGAAATGCACCTGCTCTAATGTTGGGGC 410
Db 301 AAGTTGCTCGCTGGACAAACATATCTCTGACAGAAATGCACCTGCTCTAATGTTGGGGC 360
Qy 411 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTCTCGA 470
Db 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTCTCGA 420
Qy 471 CTTCCAACTCGAATAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 530
Db 421 CTTCCAACTCGAATAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 531 ATAGAACTGGCTGTCGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 590
Db 481 ATAGAACTGGCTGTCGCGGCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACAGTT 540
Qy 591 TACAGATCCCCATACACAGTGAACAGAGGAGGAAAGCAACCAACCATGGCAGAGTCCA 650
Db 541 TACAGATCCCCATACACAGTGAACAGAGGAGGAAAGCAACCAACCATGGCAGAGTCCA 600
Qy 651 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 710
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 711 GAGCCACACCTTCCATGCTGTAGCCAGAGAGAGGGGTGAGGAGTCTTCCCTGTGTC 770
Db 661 GAGCCACACCTTCCATGCTGTAGCCAGAGAGAGGGGTGAGGAGTCTTCCCTGTGTC 720
Qy 771 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGAACTTCTTCTGCTCTCAAGCAAG 830
Db 721 GTAGCTTTTCACTGTAAAGCTTCACTTAAAGAGAGAACTTCTTCTGCTCTCAAGCAAG 780
Qy 831 GAGATGGGCTCCCAAGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 890
Db 781 GAGATGGGCTCCCAAGTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGAC 840
Qy 891 GGGAAAGCATCACTCATGAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 950
Db 841 GGGAAAGCATCACTCATGAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
```

Db 841 GGGAAAGCATCACTCATGAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
Qy 951 GATCTGTGTGCTGTTTTTGTGTGTGTAGAAATGTCAGATGAAGAGTTTCAATTCACCCCATC 1010
Db 901 GATCTGTGTGCTGTTTTTGTGTGTGTAGAAATGTCAGATGAAGAGTTTCAATTCACCCCATC 960
Qy 1011 TGTGAAGATGCCACTTTTCAGAGGTACCAAGGAAGGAGAGATGCCCGTGGCTTGGTG 1070
Db 961 TGTGAAGATGCCACTTTTCAGAGGTACCAAGGAAGGAGAGATGCCCGTGGCTTGGTG 1020
Qy 1071 GTTTCATGTGGCCCCAGCATCTGTGTGTGTGGACAGAGGTACCAAGAGGAGG 1130
Db 1021 GTTTCATGTGGCCCCAGCATCTGTGTGTGTGGACAGAGGTACCAAGAGGAGG 1080
Qy 1131 TTTGGGCTGTACACCCAGCACTTGTGTCTGTGAATGAGAACTGTGTCTGTAGTTTCAACACCTT 1190
Db 1081 TTTGGGCTGTACACCCAGCACTTGTGTCTGTGAATGAGAACTGTGTCTGTAGTTTCAACACCTT 1140
Qy 1191 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTTCATCCACCCGGACATCTTCCCTCTGTC 1250
Db 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTTCATCCACCCGGACATCTTCCCTCTGTC 1200
Qy 1251 ACCAGTTTCCGCTGTGAAGAGGAGGCCCCACCCCTCAGTGTGCCATGTTTCAGGGTCAA 1310
Db 1201 ACCAGTTTCCGCTGTGAAGAGGAGGCCCCACCCCTCAGTGTGCCATGTTTCAGGGTCAA 1260
Qy 1311 TGCCTCTCAAGTACAGCTCCGTCCTCCAGGAGGAGTGGCAGAGGATGCCATTTACT 1370
Db 1261 TGCCTCTCAAGTACAGCTCCGTCCTCCAGGAGGAGTGGCAGAGGATGCCATTTACT 1320
Qy 1371 TGCATCTGTAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCACGACAGCGTG 1430
Db 1321 TGCATCTGTAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCACGACAGCGTG 1380
Qy 1431 CAGGAGTACAGGAGAGTGCAGGACGCGCCAGCGCCAGCAGAGAGAAAGAGTCAGTAC 1490
Db 1381 CAGGAGTACAGGAGAGTGCAGGACGCGCCAGCGCCAGCAGAGAGAAAGAGTCAGTAC 1440
Qy 1491 CCAGAAATCATCTTCTTGGAAACAGGGTGTGCCATCCCGATGCCAGATTCGAAATTCAGT 1550
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGGTGTGCCATCCCGATGCCAGATTCGAAATTCAGT 1500
Qy 1551 GCCACACTTGTCAACATAAGCCCGACACGCTCTGCTACTTGGACTGTGTGGGACACA 1610
Db 1501 GCCACACTTGTCAACATAAGCCCGACACGCTCTGCTACTTGGACTGTGTGGGACACA 1560
Qy 1611 TTTGGGACAGCTGTGCCGTCAATTACGAGACACAGGTGAGAGGTCCTGGGACCCCTGGCT 1670
Db 1561 TTTGGGACAGCTGTGCCGTCAATTACGAGACACAGGTGAGAGGTCCTGGGACCCCTGGCT 1620
Qy 1671 GCTGTGTTGTGTCCACCTGTGACCCAGATACACACAGGGTTCGCCAAGTATCTTGTCTG 1730
Db 1621 GCTGTGTTGTGTCCACCTGTGACCCAGATACACACAGGGTTCGCCAAGTATCTTGTCTG 1680
Qy 1731 CAGAGAGACCGCTTGTGSCATCTTTGGAAAGCGCTTACCCCTTGTGTGTGTGCTGCC 1790
Db 1681 CAGAGAGACCGCTTGTGSCATCTTTGGAAAGCGCTTACCCCTTGTGTGTGTGCTGCC 1740
Qy 1791 CCCAACCCAGCTCAAAAGCCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGATCCTGCAC 1850
Db 1741 CCCAACCCAGCTCAAAAGCCTGTGCTCCAGCAGTACCAACACAGTGCAGAGGATCCTGCAC 1800
Qy 1851 CACATCAGTATGATTCCTGCGCAAAATGCTTCAGGAAGGGGTGAGATCTCAGTCTCTGCA 1910
Db 1801 CACATCAGTATGATTCCTGCGCAAAATGCTTCAGGAAGGGGTGAGATCTCAGTCTCTGCA 1860
Qy 1911 GTGGAAGATTGATCAGTTCGCTGTGCGAATCATGTGATTTGGAGAGTTTCAGACCTGT 1970
Db 1861 GTGGAAGATTGATCAGTTCGCTGTGCGAATCATGTGATTTGGAGAGTTTCAGACCTGT 1920
Qy 1971 CTGCTGCGGCACCTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACACCTCTGGCTGGAAA 2030
Db 1921 CTGCTGCGGCACCTGCAAGCATGCGTTTGGCTGTGCGCTGGTGACACACCTCTGGCTGGAAA 1980

Qy 2031 GTGGTCTATTTCCGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAGATGCC 2090
Db 1981 GTGGTCTATTTCCGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAGATGCC 2040
Qy 2091 ACCCTCTGATACATGAAGCCACCTTGGAAAGATGGTTTGGAAAGGAGCAGTGGAAAAG 2150
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGATGGTTTGGAAAGGAGCAGTGGAAAAG 2100
Qy 2151 ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATCGGATGAACGCGGAGTTCAAT 2210
Db 2101 ACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATCGGATGAACGCGGAGTTCAAT 2160
Qy 2211 ATGCTGAACACATTCAGCAGCGCTATGCCAAGTCCCTCTTCAGGCCCACTTCAGC 2270
Db 2161 ATGCTGAACACATTCAGCAGCGCTATGCCAAGTCCCTCTTCAGGCCCACTTCAGC 2220
Qy 2271 GAGAAAGTGGGAGTTGCTTTTGACCATGAAGTCTGTCTTGGAGACTTTTCCAAACAATG 2330
Db 2221 GAGAAAGTGGGAGTTGCTTTTGACCATGAAGTCTGTCTTGGAGACTTTTCCAAACAATG 2280
Qy 2331 CCCAAGCTGATTTCCCACTGAAAAGCCTGTTTGTGTGGCAGCATCGAGGAGTGGAGAG 2390
Db 2281 CCCAAGCTGATTTCCCACTGAAAAGCCTGTTTGTGTGGCAGCATCGAGGAGTGGAGAG 2340
Qy 2391 CGCAGGAGAGGAGGAGTCCGCGAGGTGCGGCGGCGCTCTGTCTCCAGGAGCTGGCA 2450
Db 2341 CGCAGGAGAGGAGGAGTCCGCGAGGTGCGGCGGCGCTCTGTCTCCAGGAGCTGGCA 2400
Qy 2451 GCGGCGCTGAGGATGGGAGCCTCAGCAAGACGGGCGCCACACAGAGGAGCCACAGGCC 2510
Db 2401 GCGGCGCTGAGGATGGGAGCCTCAGCAAGACGGGCGCCACACAGAGGAGCCACAGGCC 2460
Qy 2511 AAGAAGTCCAGAGCCCACTGCA 2531
Db 2461 AAGAAGTCCAGAGCCCACTGCA 2481

RESULT 8
US-09-988-687-1
; Sequence 1, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988, 687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-1

Query Match 83.9%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1279 GACTGCAATACATGATGATTCATAGCTAGGCGCTTGAGGCTCCCGAGTTCCTCAGGAGAGT 1338
QY 1428 GTGAGAGAGTACAGGAGAGTGGCGAGGAGCGGCCCGCCAGCAGAGAGAGAGAGTACAG 1487
Db 1339 GTGGAGAGATATCGGAAGAACGTCGAGGAGAACCCAGCGCCAGCAGAGAGAGAGAGCAG 1398
QY 1488 TACCCAGAGATCATCTCTCTTGGAAAGAGGTCTGCCATCCGATGAAGATTGGAATGTG 1547
Db 1399 TATCTCGAAATGTCCTCTCTGGGTAGGGGTCTGCCATCCCAATGGAGATCCGAATGTG 1458
QY 1548 AGTGCCACACTCTGCAACATAAGCCCGACACGCTCTCTGCTACTGCACTGTGTGAGGCG 1607
Db 1459 AGTTCACACTCTGCAACCTAGCCCTGACAGTCTGCTGATGTGGAGAGGCG 1518
QY 1608 ACATTTGGGAGGTGTCGCGTCTATACGAGACACAGGTGACAGGCTCTCGGACCCCTG 1667
Db 1519 ACTTTTGGGAGGTGTCGCGTCTATACGAGACACAAATAGACCGAGTCTATGACGCTC 1578
QY 1668 GCTGCTGTGTGTGTCACCTGACGAGCAGATCACACACGCGGTTCGCAAGTATCTTG 1727
Db 1579 ACGGCTGTGTGTGTCACCTGACGCGGACACACACGCGGTTCGCAAGTATCTTG 1638
QY 1728 CTGCAGAGAGAACCGCTTGGCATCTTTGGGAAAGCGCTTCACCTTTGCTGTGTT 1787
Db 1639 CTGCAGAGAGAGATGCTGTGGCATCTCTGGGAAAGCGCTTCAGCGCTTGTGTGGTG 1698
QY 1788 GCGCCCAACAGCTCAAGCGTGGCTCCAGCAGTACACAAACAGTGCAGGAGGTCTG 1847
Db 1699 GCTCTACCGAGCTCAGGCGCTGGCTGCAGCAGTATCACAACTGCAGGAGATCTG 1758
QY 1848 CACCATCAGTATGATTCGTGCAAAATGCTTCAGGAAGGGGCTGAGATCTCCAGTCT 1907
Db 1759 CACCATCAGTATGATTCGTGCAAAATGCTTCAGGAAGGGGAGGAGGTCTCCAACT 1818
QY 1908 GCAGTGGAAAGATGATCAGTTCGCTGTTGCGAACATGATTTGGAAGAGTTCAGACC 1967
Db 1819 ACATTTGGAAGGCTGATAGTCTGTTGGAAACATGATGATTAAGAAATTTTCAGACC 1878
QY 1968 TGTCTGCTGCGGACATGCAAGCATGCTTTGGCTGCGCTGCTGTCACACCTCTGCGTGG 2027
Db 1879 TGCTGTGACGCACTGCAAGCATGCTTTTGCTGTGCATCTGATCATCTGCTGGTGG 1938
QY 2028 AAAGTGTCTATTTCCGGGACACCATGCTGCGTGGAGGCTCTGCTGCGGATGGGAAAGAT 2087
Db 1939 AAAGTGTCTACTCGGGGATACCATGCTGAGGCTCTGCTGCGATGGGAAAGAT 1998
QY 2088 GGCACCTCTGATATACATGAAGCCACCTGGAAGATGTTTGGAGAGAGACAGTGGAA 2147
Db 1999 GGCACCTCTGATATCATGAAGCCACTGGAAGATCTGGAAGAGAGAGAGATGAG 2058
QY 2148 AAGACACAGACACAACTGCCAAGCCATCAGCGTGGGATCGGATGAACGCGAGTTTC 2207
Db 2059 AGGACACAGACACCACTCCCGAGCTATTAATGTGGGATCGGATGAATGGGAGTTTC 2118
QY 2208 ATTAAGTGAACACTTCAGCAGCGCTATGCAAGGTCCCGCTCTTCAGCGCCCACTTC 2267
Db 2119 ATCATGCTGAACCACTTCAGTACGCGGTACGCAAGATCCCGCTTTCAGCGCTGACTTC 2178
QY 2268 AGCGAAGAGTGGAGTTCGCTTGGACCATGAAGGTCTGCTTGGAGACTTTCACACA 2327
Db 2179 AAGGAAAGTGGCATCGCTTGGACCATGAAGGTCTGNTTGGAGACTTTCGCGACA 2238
QY 2328 ATGCCCAAGCTGATTCGCCCACTGAAAGCCCTGTTTGTCTGGGACATCAGGAGATGGAG 2387
Db 2239 GTGCCCAAGCTGATTCGCCCACTGAAGGCCCTGTTTGGAGGTGACATTTGAAGAGATGGTG 2298
QY 2388 GAGCGAGGAGAGAGCGGAGTTCGCGAGGTGCGGCGCGCCCTCTCTGCTCCAGGAGCTG 2447
Db 2299 GAAAGCGAGGAGAGAGGAGGTACGGCTGGTGGAGAGAGCGCCCTCTCTGACC--CAGCAG 2355
QY 2448 GCAGGCGGCTTGAGGATGGGAGCTCTAGCAGAGCGGAGGAGGAGGAGGAGGAGGAG 2506
Db 2356 GCAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2414

RESULT 10

US-09-988-687-221
; Sequence 221, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687

; PRIORITY FILING DATE: 2001-11-20

; PRIORITY APPLICATION NUMBER: 09/564,805

; PRIORITY FILING DATE: 2000-05-05

; PRIORITY APPLICATION NUMBER: US 60/107,468

; PRIORITY FILING DATE: 1998-11-06

; PRIORITY APPLICATION NUMBER: 09/434,382

; PRIORITY FILING DATE: 1999-11-05

; NUMBER OF SEQ ID NOS: 240

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 221

; LENGTH: 2470

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2466)

US-09-988-687-221

Query Match 55.6%; Score 1645.6; DB 9; Length 2470;

Best Local Similarity 81.6%; Pred. No. 0;

Matches 1958; Conservative 0; Mismatches 417; Indels 24; Gaps 4;

QY 108 CGCACCATATCGCAGGACCCCGCCGCGAGCGGCGCGCAAGGACCCGCTGGGCGAC 167

Db 40 CGCACCATATCGCAGGAGTTCGGCTCGTGGCGGCGCCACCCAAAGACCCACTGGCAGAC 99

QY 168 CTGCGCACGCGAGAGAGCGGACCGTCTGGGGTCTCCGGCGGCCCAAAACCCGCTGAC 227

Db 100 CTGCGTACGCGGAGAGCGCGGCC-----GGGTCCCGGGGGCCGAAACCCGCTGAC 153

QY 228 CTGCAAGTGGTGGCAGCGGCTAGCGGAGCTCGGCGCGCGGCTCTAGCTTCTCCGAG 287

Db 154 CTGCAAGTGGTGGCGGCGGCGCGGAGCGCGGCGGCTCTCTATGCTTCTTCGAA 213

QY 288 TTCAACCGGTATCTCTCAACTGTGAGAGCGGCTTCAGAGACTCATCAGGAGCACAAG 347

Db 214 TACAACAGGTACCTTTTAACTGCGGAGAGCGGTCCACGACTTATGAGGAAACACAG 273

QY 348 TTAAGAGTTCGCTCGCTGGCAACAATATTCCTGACACGAATGCACTGGTCTTAATGTTGG 407

Db 274 ACTGAAGTTCGCTCGCTGGCAACAATATTCCTGACCTCGATGCTGTCAAATGTTGG 333

QY 408 GGTCTAAGTGAATGATTTCTTTAAAGAAACCGGGCTTCAAAGTGTGATCTTTCT 467

Db 334 GGTCTGTTGGAATGATTTTAACTTTAAAGAAACCGGGCTTCCCAATGTTGTTCTGTCT 393

QY 468 GGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGGTCCATTGAA 527

Db 394 GGACCAACACAGCTGGGAAATATCTAGAACAATCAAAATATTTCTGGTCCATTGAA 453

QY 528 GGAATAGAACTGGTGTGCGGCGCCCACTCTGCGCCAGAAATACGAGGATGAACACATGACA 587

Db 454 GGAATAGAACTGGCGGTGCGGCTCACTCTGCAACAGAAATACAGGATGAGACCATGACT 513

QY 588 GTTTACCATGATCCCATACACAGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 647

Db 514 GTTTACCATGATCCCATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 573

```
Qy 648 CCAGAAAGCCCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGATCCGAGTCGAATGAA 707
Db 574 CCCAGAAATCTCCCAACAGGCTCAGTCCCAACAGTTCATCGGATCTGGATCAGCTGAA 633
Qy 708 AATGAGCCACACTTCCACATCGGTGTAGCCAGAGAGAGGGGTCAGGAGCTCTTCCCTG 767
Db 634 AATGGGC-----AGTCCACAGGAAGCATGGGGCAGGGAC-CCTCCTTA 678
Qy 768 GTCTGAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCAAAGCA 827
Db 679 GTGGTAGCTTTTGTCTGCAAGCTTCACTTGAGGAAGGAATACTTCTGGTGTAAAGCA 738
Qy 828 AAGGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATATTCTGCTGTCAAG 887
Db 739 AAGGAGCTGGGCTTCTCTGTTGGGACGGCCGCTTGTGACCCCATATTGCTGTCAAG 798
Qy 888 GACGGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCT 947
Db 799 GACGGGAAGAGTATCACTTACGAAGGAAGAGAGATTGCTGTGAAGAGCTTGTACACCC 858
Qy 948 CAGATCTCTGTGCTGCTTTTGTGTGTGTAGAAATGTCAGATGAAAGCTTCATTCAACCC 1007
Db 859 CCAGATCTCTGTGCTGTATTCATCTGTGTGTAGAGTGTCTGTATGAAGGATTCATCTGCC 918
Qy 1008 ATCTGTGAGATGCCACCTTTTCAAGAGTACCAAGAAAGGAGATGCCCGTGGCCCTTG 1067
Db 919 ATCTGTGAGAACGACACCTTTTAAAGGTACCAAGGAGAGGCTGTATGCACTGTGGCGCTG 978
Qy 1068 GTGGTTTCATGTGGCCCGCCAGCATCTGTCTGTGGACAGCAGGTACAGCAGTGGATGGAG 1127
Db 979 GTGGTCCACATAGCCCGCAGATCTGTACTATCGACAGCAGATACACAGTGGATGGAG 1038
Qy 1128 AGTTTGGGCTGTGACACCCAGCATTGTGCTGTGAATGAGAACTGTGCTCAGATTCAACAAC 1187
Db 1039 AGTTTGGGCTGTGACACAGCAGCATTGTGCTGTGAATGAGAACTGTGCTCAGATTCAACAAC 1098
Qy 1188 CTTCCAGGCCACAAATTCAAACCCAGCTCAACTCATCCACCGGACATCTTCCCTG 1247
Db 1099 CTGCGCAGCCACAAATTCAGACCCAGCTCAGCCTCATCCACCCCTGACATCTTCCCCCGAG 1158
Qy 1248 CTCACAGTTTCCGCTGTAAGAAAGAGGGCCGCCACCTCAGTGTGCCATGTTTCAGAGT 1307
Db 1159 CTTACAGCTTCTATAGTAAGAGAGAGGGTCCACCTCAGGCTGCCAACAGTTTCGGGT 1218
Qy 1308 GAATGCCCTCTCAAGTACAGCTCCGTCCAGAGGGAGTGGCAGAGGATGCCAATTATT 1367
Db 1219 GAATGCCCTCTCAAGTATTTCAGTCCGCCCCCAAGAGAGAGTGGCAGAGGATACCAACTC 1278
Qy 1368 ACTTGCATCTGAGGAATTCATAGTTGAGGCGCTGCGAGCTTCCCAACTTCCAGCAGAC 1427
Db 1279 GACTCAATACTGATGAATTCATAGCTAGGCTTGGAGCTCCCCAGTTTTCAGGAGAT 1338
Qy 1428 GTGAGGAGTACAGGAGGAGTGCAGAGAGGGCCCGAGCCCGCAGCAGAGAGAAAGAGTCAG 1487
Db 1339 GTGAGGAGTATCGAGAGAACTGTCAGAGAAACCCAGACCCCGCAGCAGAGAAAGAGCCAG 1398
Qy 1488 TACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGCAAAATGTC 1547
Db 1399 TATCTGAAATTTGCTTCTTGGGTACGGGTCTGCCATCCCAATGGAGATCCGAAATGTC 1458
Qy 1548 AGTGCCACACTTGTCAACATAGCCCGCAGCAGCTCTCTGCTACTGACTGTGTGGTAGGGC 1607
Db 1459 AGTTCCACACTGTCAACTAAGCCCTGACAAAGTCAGTGTCTCTGGATTGTGGAGAGGC 1518
Qy 1608 ACATTTGGCAGCTGTGCGCTCATTCAGGAGACCAAGGTGGACAGGCTCTGGGCAACCTG 1667
Db 1519 ACTTTTGGCAGTTGTGGCTCATTCAGGACAGCAAAATAGACCCAGTCTTATGACGCTC 1578
Qy 1668 GCTGCTGTGTTGTCTCCACTGTGACGCGAGATCAACCAACGAGGCTTCCCAAGTATCTTG 1727
Db 1579 ACGGCTGTGTTGTCTCCACCTGTCACGCCGACCCACCAACGAGGCTTCTGTAATATCTTG 1638
```

```
Qy 1728 CTGCAGAGAGAACCGCGCTTGGCATCTTTGGRAAGCGCTTACCCCTTTCCTGCTGCTGTT 1787
Db 1639 CTGCAGAGAGAGCATGCTTGGCATCTCTGGGAAACCCCTTCCAGCCCTTCTTGTGTG 1698
Qy 1788 GCCCCCAACAGCTCAAAAGCCTGGCTCCAGCAGTACCACAAACAGTGCAGGAGGCTCTG 1847
Db 1699 GCTCTACCCAGCTCAGGCGCTGGCTGCAGCAGTATCACAACCACTGCCAGGAGATCTG 1758
Qy 1848 CACCATCATGATATGATTCCTGCAAAATGCTTCCAGAAAGGGCTGAGATCTCCAGTCT 1907
Db 1759 CACCACGCTCAGTATGATTCCTGCAAAATGCTTCCAGAAAGGGCAGAGGTCCTCAATACT 1818
Qy 1908 GCAGTGAAGATTTGATCAGTTCGCTGTGCAACATGTGATTTGGAAAGATTTTCAGACC 1967
Db 1819 ACATTTGAAGAGGCTGATAAGCTTGTGTGAAACATGTGACTTAGAAGAAATTTTCAGACC 1878
Qy 1968 TGTCTGTGTGGGCACTGCAAGCATGCTTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGG 2027
Db 1879 TGCCTGTGTACGGCACTGCAAGCATGCTTTTGGCTGTGCACTGGTACATTCATCTGGCTGG 1938
Qy 2028 AAAGTGTCTTATTCGGGGACACCATGCCCTGCGAGGCTCTGTGCTCGGATGGGAAAGAT 2087
Db 1939 AAAGTGTCTTACTCGGGGATACCATGCCCTGTGAGGCTCTGGTCCAGATGGGAAAGAT 1998
Qy 2088 GCCACCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGAGTGGAA 2147
Db 1999 GCCACCTCTGATACATGAAGCCACCTCTGAGGATCNCITTTGAAGAGAGAGAGTAGAG 2058
Qy 2148 AAGACACACAGCACAACGCTCCAAAGCCATCAGCGTGGGGATGCGGATCAACCGGAGTTC 2207
Db 2059 AGGACACACAGCACCACCTCCAGGCTATTAAATGTGGGGATGCGGATCAATGCGGAGTTC 2118
Qy 2208 ATTATGTGAACCACTTTCAGCAGCGCTATGCAAGGCTCCCTCTTTCAGGCCCAACTTC 2267
Db 2119 ATCATGTGAACCACTTTCAGTCAAGGCTACGCNAAGATCCCTCTTTCAGGCCCTGACTTC 2178
Qy 2268 AGCGAGAAAGTGGGAGTTCCTTTGACCATGAAGCTCTGCTTTGGAGACTTTTCCAACA 2327
Db 2179 AACGAGAAAGTGGGATTCGCTTTGACCATGAAGTCTGNTTTGGAGACTTCCCGACA 2238
Qy 2328 ATGCCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTTGTGCGCACATCGAGGAGATGGAG 2387
Db 2239 GTGCCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTTGCAGGTGACATTGAAGAGATGGTG 2298
Qy 2388 GAGCGCAGGAGAAAGCGGAGCTGCGGAGAGTGGCGGCGGCCCTCTCTGTCAGGGAGCTG 2447
Db 2299 GAACGCGAGGAGAAAGAGGAGCTACGGCTGTGTCGAGCAGCCCTCTCTGACC---CAGCAG 2355
Qy 2448 GCAGCGCGCTTGGAGATGGGAGCCTCAGCAGAAAGCGGGCCCAACAGAGGAGCCACA 2506
Db 2356 GCAGCAGCCAGAGGACAGAGAAACCCCAACAGAAAGCGGGCCCAACAGATGAACACACA 2414
```

RESULT 11

US-09-833-381-2039

; Sequence 2039, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 2039

; LENGTH: 783

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature


```
; LOCATION: (1)....(783)
; OTHER INFORMATION: n = A,T,C or G
; US-09-833-381-2039

Query Match      24.8%; Score 734.8; DB 10; Length 783;
Best Local Similarity 98.3%; Pred. No. 6.7e-222;
Matches 772; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

Qy 141 CGGCCGCGCAAGACCCGCTGGCGCACCTGCGACCGAGAGAGCGCGGACCGCTGGGG 200
Db 1 CGGCCGCGCAAGACCCGCTGGCGCACCTGCGACCGAGAGAGCGCGGACCGCTGGGG 60

Qy 201 TGCTCGGCGGCCCAACACCGGTGACCTGCGAGGTGGTGGCAGCGGTAGCGGACTCG 260
Db 61 TGCTCGGCGGCCCAACACCGGTGACCTGCGAGGTGGTGGCAGCGGGTA-CCGGGACTCG 119

Qy 261 GGGCGCGCGCTCTACGCTCTTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAGCG 320
Db 120 GGGCGCGCGCTCTACGCTCTTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAGCG 179

Qy 321 GTTCAGAGACTCATGCGAGGACACAAAGTTAAAGTTGCTCGCTGGACCAACATATTCCTG 380
Db 180 GTTCAGAGACTCATGCGAGGACACAAAGTTAAAGTTGCTCGCTGGACCAACATATTCCTG 239

Qy 381 ACACGAAATGCACTGCTTAATGTTGGGGCTTAAGTGGAAATGATCTTACTTTAAAGGAA 440
Db 240 ACACGAAATGCACTGCTTAATGTTGGGGCTTAAGTGGAAATGATCTTACTTTAAAGGAA 299

Qy 441 ACCGGCTTCCAAAGTGTGACTTTCTGACCTCCACAACTGGAATACTCTCGAAGCA 500
Db 300 ACCGGCTTCCAAAGTGTGACTTTCTGACCTCCACAACTGGAATACTCTCGAAGCA 359

Qy 501 ATCAAAATATTTCTGGTCCATGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCC 560
Db 360 ATCAAAATATTTCTGGTCCATGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCC 419

Qy 561 CCAGAATACAGGATGAAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGG 620
Db 420 CCAGAATACAGGATGAAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGG 479

Qy 621 AGGGGAAACCAACCATGGGAGCTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 680
Db 480 AGGGGAAACCAACCATGGGAGCTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 539

Qy 681 CGATCTTCAGACTCCGAGTTCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCA 739
Db 540 CGATCTTCAGACTCCGAGTTCGAATGAAATGAGCCACACCTTCCACATGGTGTAGCCA 599

Qy 740 GAGAGAGGGGTGAGGACTCTTCCCTGGTGTGAGCTTTTCACTGTAAAGTTCACCTAAA 799
Db 600 GAGAGAGGGGTGAGGACTCTTCCCTGGTGTGAGCTTTTCACTGTAAAGTTCACCTAAA 659

Qy 800 GAGAGAAACTCTTGGTCTCAAGCAAGGAGATGGGCTCCCACTTGGGACAGCTCG 859
Db 660 GAGAGAAACTCTTGGTCTCAAGCAAGGAGATGGGCTCCCACTTGGGACAGCTCG 718

Qy 860 CATCGCTCCCATCATTTGCTGTCAAGGACGGGAAAGCATCACTCATGAAAGGAGAGA 919
Db 719 CATCNCCTCCCATCATTTGCTGTCAAGGACGNCNGNAAACCAACCCATTAAGGAAAGA 778

Qy 920 GATTT 924
Db 779 GATTT 783
```

RESULT 12

```
US-09-988-626-28
; Sequence 28, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
```

```
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16494-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-626-28
```

```
Query Match      22.2%; Score 657.2; DB 9; Length 26664;
Best Local Similarity 99.5%; Pred. No. 2e-196;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2297 CATGAAGTGTCTGTTGGAGACTTTTCCAAATGCCCAGCTGATTTCCCCCACTGAAAGC 2356
Db 25805 CTTCTAGGTCTCTTTGGAGACTTTTCCAAATGCCCAGCTGATTTCCCCCACTGAAAGC 25864

Qy 2357 CTTGTTTGTGCGCAGACTCGAGGAGATGGAGAGCGCGGAGAGAGCGGAGCTCGCGCA 2416
Db 25865 CTTGTTTGTGCGCAGACTCGAGGAGATGGAGAGCGCGGAGAGAGCGGAGCTCGCGCA 25924

Qy 2417 GGTGCGGCGGCGCTCTGTTCCAGGAGCTGCGAGCGGCGCTTGGAGGATGGGAGCTCA 2476
Db 25925 GGTGCGGCGGCGCTCTGTTCCAGGAGCTGCGAGCGGCGCTTGGAGGATGGGAGCTCA 25984

Qy 2477 GCAGAAGCGGCGCCACACAGAGAGCCACAGCCCAAGAGGTTCAGAGCCAGTGAAGATC 2536
Db 25985 GCAGAAGCGGCGCCACACAGAGAGCCACAGCCCAAGAGGTTCAGAGCCAGTGAAGATC 26044

Qy 2537 TGGAGAGACCTGAACTCAGAAGCTGTGTCTTCTGCCCCACGACCGACCGTATCTG 2596
Db 26045 TGGAGAGACCTGAACTCAGAAGCTGTGTCTTCTGCCCCACGACCGACCGTATCTG 26104

Qy 2597 CCCTCTTGTGTTAGAGCTGAAGAGCAGCTCCCCAGGAGGAGCTCAGATAGGTTG 2656
Db 2597 CCCTCTTGTGTTAGAGCTGAAGAGCAGCTCCCCAGGAGGAGCTCAGATAGGTTG 2656
```

```
Db 26105 CCCTCTTGCTGGTAGAAGCTGAAGACGACGCTGCCCGAGGAGGAGCTCAGGATAGGTG 26164
Qy 2657 GTATGGAGCTGTGCGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTTA 2716
Db 26165 GTATGGAGCTGTGCGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTTA 26224
Qy 2717 GGACTGGTGCCTGGCACAGCGCGGCGGCGGAGGCTGCCACAGGAAGCAAGCAGATGAA 2776
Db 26225 GGACTGGTGCCTGGCACAGCGCGGCGGCGGAGGCTGCCACAGGAAGCAAGCAGATGAA 26284
Qy 2777 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 2836
Db 26285 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 26344
Qy 2837 TCTAATCCAGCAAGTGTATTCCTGCACACAGACAGCAAGCAGAGTAACAGATCAGTGG 2896
Db 26345 TCTAATCCAGCAAGTGTATTCCTGCACACAGACAGCAAGCAGAGTAACAGATCAGTGG 26404
Qy 2897 GTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTTGAGTTGC 2956
Db 26405 GTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTTGAGTTGC 26464
Qy 2957 AA 2958
Db 26465 AA 26466
```

RESULT 13

```
US-09-988-687-28
; Sequence 28, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIORITY FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; OTHER INFORMATION: 13032-13104;
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
```

```
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-988-687-28
```

```
Query Match 22.2%; Score 657.2; DB 9; Length 26664;
Best Local Similarity 99.5%; Pred. No. 2e-196;
Matches 659; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 2297 CATCAAGGTCTGCTTGGAGACTTTCCAAACAATGCCCAAGCTGATTTCCCCACCTGAAGC 2356
Db 25805 CTTCTAGGCTGTGCTTGGAGACTTTCCAAACAATGCCCAAGCTGATTTCCCCACCTGAAGC 25864
Qy 2357 CCTCTTTGCTGGCGACATCGAGGAGATGGAGGAGCGGAGGAGAAAGCGGAGCTGCGGCA 2416
Db 25865 CTTGTTTGTGCGACATCGAGGAGATGGAGGAGCGGAGGAGAAAGCGGAGCTGCGGCA 25924
Qy 2417 GGTGCGGCGGCGCTCTGCTCCAGGGAGCTGGCAGGCGGCTTGGAGGATGGGGAGCCCTCA 2476
Db 25925 GGTGCGGCGGCGCTCTGCTCCAGGGAGCTGGCAGGCGGCTTGGAGGATGGGGAGCCCTCA 25984
Qy 2477 GCAGAACGGGGCCACACAGAGGAGCCACAGGCCAAGAGGTACAGACCCAGTGAAGATC 2536
Db 25985 GCAGAACGGGGCCACACAGAGGAGCCACAGGCCAAGAGGTACAGACCCAGTGAAGATC 26044
Qy 2537 TGGGAGACCTGAACTCAGAAAGGCTGTGCTTCTCCCCACGACGACCCCTATCTG 2596
Db 26045 TGGGAGACCTGAACTCAGAAAGGCTGTGCTTCTCCCCACGACGACCCCTATCTG 26104
Qy 2597 CCCTCTCTTGTGTTAGAAAGCTGAAGACGACGCTCCCCCAGGAGGACGCTCAGGATAGGTG 2656
Db 26105 CCCTCTCTTGTGTTAGAAAGCTGAAGACGACGCTCCCCCAGGAGGACGCTCAGGATAGGTG 26164
Qy 2657 GTATGGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTA 2716
Db 26165 GTATGGAGCTGTGCGGAGGCTTGGGCTCCACATGAAGCACTAGTCTATAGATGCTCTTA 26224
Qy 2717 GGACTGTGCTGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAGAGATGAA 2776
Db 26225 GGACTGTGCTGGCACAGCGCGGCGGCGGAGGCTGCCACACGGAAGCAGAGATGAA 26284
Qy 2777 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 2836
Db 26285 CTAATTTTCATTTCAAGGAGTGTGTTTAAAGAGTCTTGGAAACAGACGGCGGCACCTTTCC 26344
Qy 2837 TCTAATCCAGCAAGTGTATTCCTGCACACAGACAGCAAGCAGAGTAACAGGATCAGTGG 2896
Db 26345 TCTAATCCAGCAAGTGTATTCCTGCACACAGACAGCAAGCAGAGTAACAGGATCAGTGG 26404
Qy 2897 GTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTTGAGTTGC 2956
Db 26405 GTCTAAGTGTCCGAGACTTAACGAAATAGTATTTTCAGCTGCAATAAAGATTTGAGTTGC 26464
Qy 2957 AA 2958
Db 26465 AA 26466
```

RESULT 14

```
US-09-988-626-27
; Sequence 27, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
```


;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/988,626
;; CURRENT FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: 09/564,805
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 27
;; LENGTH: 655
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(228)
;; OTHER INFORMATION: exon 24
;; NAME/KEY: polyA signal
;; LOCATION: (636)..(641)
US-09-988-626-27

Query Match 22.1%; Score 655; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2304 GTCTGTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTT 2363
Db 1 GTCTGTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTT 60
QY 2364 GTTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGGCGAGGTGCGG 2423
Db 61 GTTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGGCGAGGTGCGG 120
QY 2424 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGTGAGATGGGAGGCTCAGCAGAG 2483
Db 121 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGTGAGATGGGAGGCTCAGCAGAG 180
QY 2484 CGGGCCCAACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAG 2543
Db 181 CGGGCCCAACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAG 240
QY 2544 CCTGAACCTCAGAAAGCTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 2603
Db 241 CCTGAACCTCAGAAAGCTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 300
QY 2604 TGTGTAGAGCTGAAGAGCGTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 2663
Db 301 TGTGTAGAGCTGAAGAGCGTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 360
QY 2664 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 2723
Db 361 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 420
QY 2724 TGCTGGCAGCGCGGCGCAGGAGCTGCCACAGCAAGCAAGCAGATGAATTAATTT 2783
Db 421 TGCTGGCAGCGCGGCGCAGGAGCTGCCACAGCAAGCAAGCAGATGAATTAATTT 480
QY 2784 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 2843
Db 481 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 540
QY 2844 CAGCAAGTGTATTCCTGACACACAGCAGACAGCAGTAAACAGATCAGTGGCTCTAAG 2903
Db 541 CAGCAAGTGTATTCCTGACACACAGCAGACAGCAGTAAACAGATCAGTGGCTCTAAG 600
QY 2904 TGTCCGAGACTTAAACGAAATAGTATTTTGTGCTGCAATTAAGATTTCAGTTTGCAG 2958
Db 601 TGTCCGAGACTTAAACGAAATAGTATTTTGTGCTGCAATTAAGATTTCAGTTTGCAG 655

RESULT 15
US-09-988-687-27
;; Sequence 27, Application US/09988687
;; Publication No. US20030045704A1
;; GENERAL INFORMATION:
;; APPLICANT: Tavtigian, Sean V.
;; APPLICANT: Teng, David H.F.
;; APPLICANT: Simard, Jacques
;; APPLICANT: Rommens, Johanna M.
;; APPLICANT: Myriad Genetics, Inc.
;; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
;; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
;; FILE REFERENCE: 2318-258
;; CURRENT APPLICATION NUMBER: US/09/988,687
;; CURRENT FILING DATE: 2001-11-20
;; PRIOR APPLICATION NUMBER: 09/564,805
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: US 60/107,468
;; PRIOR FILING DATE: 1998-11-06
;; PRIOR APPLICATION NUMBER: 09/434,382
;; PRIOR FILING DATE: 1999-11-05
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 27
;; LENGTH: 655
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(228)
;; OTHER INFORMATION: exon 24
;; NAME/KEY: polyA signal
;; LOCATION: (636)..(641)
US-09-988-687-27

Query Match 22.1%; Score 655; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2304 GTCTGTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTT 2363
Db 1 GTCTGTTTGGAGACTTCCAAACATGCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTT 60
QY 2364 GTTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGGCGAGGTGCGG 2423
Db 61 GTTGGCGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTGGCGAGGTGCGG 120
QY 2424 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGTGAGATGGGAGGCTCAGCAGAG 2483
Db 121 GGGGCCCTCTGTCCAGGAGCTGGCAGGGGCTGTGAGATGGGAGGCTCAGCAGAG 180
QY 2484 CGGGCCCAACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAG 2543
Db 181 CGGGCCCAACAGAGGAGCCACAGGCGCAAGAGGTGAGAGCCAGTGAAGATCTGGGAG 240
QY 2544 CCTGAACCTCAGAAAGCTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 2603
Db 241 CCTGAACCTCAGAAAGCTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 300
QY 2604 TGTGTAGAGCTGAAGAGCGTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 2663
Db 301 TGTGTAGAGCTGAAGAGCGTGTGTCTTCTGCCCAACAGCAGCAGCCGATCTGCCCTCT 360
QY 2664 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 2723
Db 361 GTGTGCCAGGCTGGGCTCCACATAGCACTAGTCTATAGATGCCTCTTAGGACTGG 420
QY 2724 TGCTGGCAGCGCGGCGCAGGAGCTGCCACAGCAAGCAAGCAGATGAATTAATTT 2783
Db 421 TGCTGGCAGCGCGGCGCAGGAGCTGCCACAGCAAGCAAGCAGATGAATTAATTT 480
QY 2784 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 2843
Db 481 CATTTCAAGGAGCTTTTAAAGAGTCTTTGGAAACAGACGGGGCACCTTTCTCTAATC 655

Db 481 CATTTCAAGGCAGTTTTTAAAGAGTCTTTGAAACAGACGGCGGCACCTTTCCTCTAATC 540
Qy 2844 CAGCAAGTGATTCCTGCACACCGAGACAGCAGAGTAAACAGGATCAGTGGGTCTAAG 2903
Db 541 CAGCAAGTGATTCCTGCACACCGAGACAGCAGAGTAAACAGGATCAGTGGGTCTAAG 600
Qy 2904 TGTCCGAGACTTAAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTGCAA 2958
Db 601 TGTCCGAGACTTAAACGAAATAGTATTTTCAGCTGCAATAAAGATTGAGTTTGCAA 655

Search completed: May 18, 2003, 02:19:09
Job time : 522 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 17, 2003, 15:56:54 ; Search time 6428.12 Seconds
(without alignments)
11232.524 Million cell updates/sec

Title: US-09-434-382-1
Perfect score: 2481
Sequence: 1 atgtggcgcttctgcgtc.....agaaggtcagcccaagtga 2481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2481	100.0	2908	9	AF304370	AF304370 Homo sapi
2	2481	100.0	2992	6	AX405825	AX405825 Sequence
3	2477.8	99.9	2997	9	BC001939	BC001939 Homo sapi
4	2477.8	99.9	3006	9	BC004158	BC004158 Homo sapi
5	2455.4	99.0	2908	9	AF308698	AF308698 Pan trogl
6	2448.2	98.7	2976	9	AK001392	AK001392 Homo sapi
7	2442.6	98.5	2893	9	AF308694	AF308694 Gorilla g
8	2349	94.7	2907	9	AK094687	AK094687 Homo sapi
9	1745	70.3	2734	9	AK094333	AK094333 Homo sapi
10	1715.2	69.1	2163	9	AK074244	AK074244 Homo sapi
11	1686.2	68.0	2712	10	AF308696	AF308696 Mus muscu
12	823	33.2	2793	9	AK094012	AK094012 Homo sapi
13	247.4	10.0	118788	9	AC005277	AC005277 Homo sapi
14	156.8	6.3	679	9	HS329331	HS329331 Homo sapi
15	146.2	5.9	740	9	AF304371S2	AF304369 Homo sapi
16	144.4	5.8	34593	10	AF348157	AF348157 Mus muscu
17	144.4	5.8	196566	10	AL663045	AL663045 Mouse DNA
C 18	135.8	5.5	167336	2	AC128361	AC128361 Rattus no
C 19	135.8	5.5	186836	2	AC097911	AC097911 Rattus no
C 20	133.4	5.4	2690	3	AF215894	AF215894 Drosophil
21	133.4	5.4	2728	3	AV119279	AV119279 Drosophil
22	133.4	5.4	115873	2	AC017383	AC017383 Drosophil
C 23	133.4	5.4	175681	3	AC007417	AC007417 Drosophil
24	133.4	5.4	190574	2	AC007352	AC007352 Drosophil
C 25	133.4	5.4	240481	3	AE003830	AE003830 Drosophil
C 26	132.4	5.3	186836	2	AC097911	AC097911 Rattus no
27	124	5.0	721	9	AF304371S1	AF304371 Homo sapi
28	115.2	4.6	167336	2	AC128361	AC128361 Rattus no
29	106.6	4.3	2704	8	AY117303	AY117303 Arabidops
C 30	106.6	4.3	2818	8	AY070061	AY070061 Arabidops
C 31	95.4	3.8	752	9	HS329367	HS329367 Homo sapi
32	90.2	3.6	1617	8	AY086950	AY086950 Arabidops
C 33	84.2	3.4	31201	8	SPAC1D4	Z69239 S pombe chr
C 34	71.8	2.9	46335	8	AC022354	AC022354 Arabidops
C 35	70	2.8	19158	8	AF188714	AF188714 Emericell
C 36	67	2.7	32329	8	SPBC3D6	Z95620 S.pombe chr
C 37	65	2.6	143209	8	AF001550	AF001550 Oryza sat
C 38	63.4	2.6	125020	9	AF429315	AF429315 Homo sapi
39	59.4	2.4	256050	1	AL627274	AL627274 Salmonell
40	58.8	2.4	75289	8	AB023046	AB023046 Arabidops
41	58.4	2.4	2517	6	AR137565	AR137565 Sequence
42	58.4	2.4	2517	6	AR139071	AR139071 Sequence
C 43	58.4	2.4	3685	8	SCYKR079C	Z28304 S.cerevisia
44	57.8	2.3	23630	1	AE008803	AE008803 Salmonell
45	56.8	2.3	23533	1	AE001039	AE001039 Archaeogl

ALIGNMENTS

RESULT 1
AF304370
LOCUS
DEFINITION Homo sapiens putative prostate cancer susceptibility protein
HPC2/ELAC2 mRNA, complete cds.
ACCESSION AF304370
VERSION AF304370.1 GI:10880932
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2908)
Tavtigian,S.V., Simard,J., Teng,D.H.F., Abtin,V., Baumgard,M.,
Beck,A., Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P.,

Desrochers, M., Dumont, M., Farnham, J. M., Frank, D., Frye, C.,
Chaffari, S., Gupta, J. S., Hu, R., Iliev, D., Janeski, T., Kort, E. N.,
Laitly, K. E., Leavitt, A., Leblanc, G., McArthur-Morrison, J.,
Pederson, A., Penn, B., Peterson, K. T., Reid, J. E., Richards, S.,
Schroeder, M., Smith, R., Snyder, S. C., Swedlund, B., Swensen, J.,
Thomas, A., Tranchant, M., Woodland, A. M., Labrie, F., Skolnick, M. H.,
Neuhausen, S., Rommens, J. and Cannon-Albright, L. A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
11175785
2 (bases 1 to 2908)
Tavtigian, S. V., Simard, J., Teng, D. H. F., Baumgard, M., Beck, A.,
Camp, N. J., Carillo, A. R., Chen, Y., Dayananth, P., Desrochers, M.,
Dumont, M., Farnham, J. M., Frank, D., Frye, C., Ghaffari, S.,
Gupta, J. S., Hu, R., Iliev, D., Janeski, T., Kort, E. N., Laitly, K. E.,
Leavitt, A., Leblanc, G., McArthur-Morrison, J., Pederson, A., Penn, B.,
Peterson, K. T., Reid, J. E., Richards, S., Schroeder, M., Smith, R.,
Snyder, S. C., Swedlund, B., Swensen, J., Thomas, A., Tranchant, M.,
Woodland, A. M., Labrie, F., Skolnick, M. H., Neuhausen, S., Rommens, J.
and Cannon-Albright, L. A.
Direct Submission
Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakarusa Way, Salt
Lake City, UT 84108, USA
Location/Qualifiers
1..2908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p"
1..2481
/codon_start=1
/product="putative prostate cancer susceptibility protein
HPC2/ELAC2"
/protein_id="AAG24441.1"
/db_xref="GI:10880933"
/translation="MWALCSLLRSAAAGTWSQRTISQAPARRPRPKDLRLHRTRE
KRFGSGGNTVYLVQVAAGSDGAALYVSEFNRYLFCNGEGVQRLWQERLKV
ARLNIIFLTRHNSVNGSLMILTRETGLPKVLSPPQLEKYLBAIKIFSGPLKG
IELAVRPHSAPEYDEMTVYQIPHSEORRGHPQSPERPLSRSDSDSEN
ENEPHLPQSVQRGVRDSSLVAFICKLHLKRGNFVLVLRKAKMGLPVGTAAPITA
AVDKGKSTHEGREILABELCTPPDPGAAVFWVECPDESFIQPCENATFORYOGKAD
APVALVHMAPASVLDVSRVQOMMERFGPDQHLVLRNENCAVHNLSFKIQTQNLNI
HPDIFPLLTFRCKKEGPTLSVPMVQSECLLKYLQRLRRWRQDRLITCPREFIVEA
LQLPNFQSVQVEYRSADQSPAPAKRSQYPIIFLGTGSAIPMKIRNVSATLVNISP
DTSLLDCEGTEFGQLCRHYGDQVRLVGLTAAVFSVHLHADHTGTPSLILQERAL
ASLQKPLHLVAVPNQKWLQOYHNCQCEVLHHISMIPAKCLQEGAEISSPAVERL
ISSLRRTCDLEEFOTCLVRHCKHAFGCALVHTSGWKVYSGDTWPCALVRMGKDAFL
LIHEATLEDGLEEAEVETKSTTSQALSVGMRMNAEFIMLNHFSQRYAKVPLSPNFS
EKUGVAFDHKKVCGDPFPTPKLIPPLKALFAGDIEBERERREKRELRQVRAALLSRE
LAGGLEDEGPQQRKATEBEPQAKKVRQA"

BASE COUNT 702 a 793 c 826 g 587 t

ORIGIN

Query Match 100.0%; Score 2481; DB 9; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGTGGGGCTTTGCTCGTGTGCGGTCGCGGCGCGGACGACCATGTCGCGAGGAGCC 60
Db 1 ATGTGGGGCTTTGCTCGTGTGCGGTCGCGGCGCGGACGACCATGTCGCGAGGAGCC 60
Oy 61 ACCATATCGAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 ACCATATCGAGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Oy 121 CGCAGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 CGCAGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Oy 181 CAGGTGTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 CAGGTGTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

QY 1321 TGCATCTCTGAGGAATTCATAGTTGAGGGCTGCGAGCTTCCCACTTCACGAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGGCTGCGAGCTTCCCACTTCACGAGAGCGTG 1380
QY 1381 CAGGAGTACAGGAGGTGCGCAGGACGCGCCAGCCCGAGCAGAGAGAAAGAGTCACTAC 1440
Db 1381 CAGGAGTACAGGAGGTGCGCAGGACGCGCCAGCCCGAGCAGAGAGAAAGAGTCACTAC 1440
QY 1441 CCAGAAATCATCTCTTGGAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCACT 1500
Db 1441 CCAGAAATCATCTCTTGGAACAGGGTCTGCCATCCCGATGAAGATTGAAATGTCACT 1500
QY 1501 GCCACACTGTCAACATAGCCCGACAGCTCTGCTACTGAGTGTGGTCAAGGCACA 1560
Db 1501 GCCACACTGTCAACATAGCCCGACAGCTCTGCTACTGAGTGTGGTCAAGGCACA 1560
QY 1561 TTTGGGCGAGCTGTGGCTCATTCAGGAGACAGGTGGACAGGGTCTGGGCACTGGCT 1620
Db 1561 TTTGGGCGAGCTGTGGCTCATTCAGGAGACAGGTGGACAGGGTCTGGGCACTGGCT 1620
QY 1621 GCTGTGTTGTCTCCACTGACGAGATACCAACAGGGTCTGCCAAGTATCTTGCTG 1680
Db 1621 GCTGTGTTGTCTCCACTGACGAGATACCAACAGGGTCTGCCAAGTATCTTGCTG 1680
QY 1681 CAGAGAGACGGCTTGGCATCTTTGGAAAGCGCTTCACTTGTGCTGTGGTGGC 1740
Db 1681 CAGAGAGACGGCTTGGCATCTTTGGAAAGCGCTTCACTTGTGCTGTGGTGGC 1740
QY 1741 CCCAACCGAGCTCAAGCGCTGCGTCCAGCAGTACCAACAGGTGGCAGAGGTCTGGAC 1800
Db 1741 CCCAACCGAGCTCAAGCGCTGCGTCCAGCAGTACCAACAGGTGGCAGAGGTCTGGAC 1800
QY 1801 CACATCAGTATGATCTCTGCCAAATGCCCTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATCTCTGCCAAATGCCCTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
QY 1861 GTGGAAGATGATCAGTTCGCTGTGCGAATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 GTGGAAGATGATCAGTTCGCTGTGCGAATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1921 CTGGTGGGCACTGAAGCATGCGTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1980
Db 1921 CTGGTGGGCACTGAAGCATGCGTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1980
QY 1981 GTGGTCTATTCCGGGACACCATGCGCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGGTCTATTCCGGGACACCATGCGCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
QY 2041 ACCCTCTGATACATGAAGCCACCTGGAAGATGTTTGGAAAGAGACGATGGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTGGAAGATGTTTGGAAAGAGACGATGGAAAG 2100
QY 2101 ACACAGACACAGCTCCCAAGCCATCAGCTGGGATGCGGATGACCGAGTTCATT 2160
Db 2101 ACACAGACACAGCTCCCAAGCCATCAGCTGGGATGCGGATGACCGAGTTCATT 2160
QY 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTCAGCCCCCACTTCAGC 2220
QY 2221 GAGAAAGTGGAGTTCCTTTGACCAATGAAGGTCTGCTTTGGAGACTTTTCCAAATG 2280
Db 2221 GAGAAAGTGGAGTTCCTTTGACCAATGAAGGTCTGCTTTGGAGACTTTTCCAAATG 2280
QY 2281 CCCAGCTGATTTCCCACTGAAGCCCTGTTGCTGGCGACATCGAGAGATGGAGAG 2340
Db 2281 CCCAGCTGATTTCCCACTGAAGCCCTGTTGCTGGCGACATCGAGAGATGGAGAG 2340
QY 2341 CGCAGGAGAAAGCGGAGTTCGCGAGGTGCGGGCGGCTCTCTGTCAGGAGTGGCA 2400
Db 2341 CGCAGGAGAAAGCGGAGTTCGCGAGGTGCGGGCGGCTCTCTGTCAGGAGTGGCA 2400
QY 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGGCCCAACACAGAGAGCCAGGCC 2460

Db 2401 GCGGCGCTGGAGATGGGAGCTTCAGCAGAAAGCGGGCCCAACACAGAGAGCCAGGCC 2460
QY 2461 AAGAAGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAGTTCAGAGCCCACTGA 2481
RESULT 2
AX405825
LOCUS AX405825 2992 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 240 from Patent WO222660.
ACCESSION AX405825
VERSION AX405825.1 GI:21439089
KEYWORDS SOURCE
ORGANISM human.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 240 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
1..2992
/organism="Homo sapiens"
/db_xref="taxon:9606"
75..2555
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34932.1"
/db_xref="GI:21439090"
/translation="MMALCSILRSAAAGRTMSQGTISOAPARRERPRKDPRLHRLTRE
KRGSGSGSPNTVYLVVAAGSDSLAALYVFSEPNRYLFCGEGVORLQWELKLV
ARLDNIETRMHNSVGLSMILTKETGLKCVLPSPQLEKYLEIKLFGSLKLG
IELAVRPHSAPEYEDETMVTYQIFHSEORRKHQWPQSPERPLSRSPSPSSSES
ENEPHLPVQSRGVRDSSLVAFIKLHLKRGFLVLKAKEMGLPVTAAIPIIA
AVKDKSITHEGREILABELCTPPDGAAFVVECPDESFIQICENATFORYOGKAD
APVALVHMAPASVLVDSRYCOOMEREGPDTQHLVLNENCASVNLRSKLTOTLNLI
HPDIFPLTSTFRCKEGPTLSVPMVQSECLLKYLQRPREWRQDAITCNBEFIVEA
LDLPNFQSQVQYRRSAQDGPAPAEKRSQYPEIIFLGTGSAIPMKIRNVSAITLVNIS
TSLLLDCGEGTGGQLRGHYGDQVDRVLGTAAVFSHLHADHTGTPSLILQERAL
ASLGLPLHLLVAPNQLKAWLQYHNQCEVLHHISMIPAKLCQEGAEISPAVERL
ISLLRTCDLEEFOTCLVRKCHAFGALVHTSGWKVYSGDTPCEALVRMGKDATL
LIHEATLEDGLEEAVEKTHSTTSQALSVGMRMNAERIMLNHFORSYAKVPLFSPNFS
EKGVAFDHMKVCFGDEPTMPKLIPLKALFAGDIEEMEERERRELRQVRAALLSRE
LAGGLEDBEPQKRAHTEEPQAKVRAQ"
BASE COUNT 725 a 807 c 859 g 601 t
ORIGIN
Query Match 100.0%; Score 2481; DB 6; Length 2992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGGGCGCTTTGCTGCTGCTGCTGCGGCGCGGAGCAGCACCATGTCGAGGAGCGC 60
Db 75 ATGTGGGCGCTTTGCTGCTGCTGCTGCGGCGCGGAGCAGCACCATGTCGAGGAGCGC 134
QY 61 ACCATATCGAGGACACCGCGCGCGGAGCGCGCGGAGGAGCAGCAGCAGCAGCAGCAGC 120
Db 135 ACCATATCGAGGACACCGCGCGCGGAGCGCGCGGAGGAGCAGCAGCAGCAGCAGCAGC 194
QY 121 CGCAGCGAGAGAGCGCGGAGCGCGGAGCGCGCGGAGGAGCAGCAGCAGCAGCAGCAGC 180
Db 195 CGCAGCGAGAGAGCGCGGAGCGCGGAGCGCGGAGGAGCAGCAGCAGCAGCAGCAGCAGC 254
QY 181 CAGGTGGTGGCAGCGGAGTGGCGGAGCTCGGCGCGCGGCTCTACGTCTTCTCCAGTTC 240
Db 255 CAGGTGGTGGCAGCGGAGTGGCGGAGCTCGGCGCGCGGCTCTACGTCTTCTCCAGTTC 314
QY 241 AACCGGTATCTTTCACTGTGTGAGAGCGGCTTCAGAGACTCATGAGGAGCACAAGTTTA 300

|||||
Db 315 AACGGTATCTCTTCAACTGTGGAGAGCGCTTCAGAGACTCATCGAGAGACAAAGTTA 374
Qy 301 AAGTTGTCTGCGCTGGACAAATATCTCTGACACAAATGCACTGTCTTAATTTGGGGGC 360
Db 375 AAGTTGTCTGCGCTGGACAAATATCTCTGACACAAATGCACTGTCTTAATTTGGGGGC 434
Qy 361 TTAAGTGAATGATTTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTCGA 420
Db 435 TTAAGTGAATGATTTCTTACTTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTCGA 494
Qy 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAAATATTTCTGTGCCATTGGAAGGA 480
Db 495 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAAATATTTCTGTGCCATTGGAAGGA 554
Qy 481 ATAGAACTGGCTGTCGGCCCCACTCTCTCCCGAGAAATACGAGGATGAAACCAATGACAGTT 540
Db 555 ATAGAACTGGCTGTCGGCCCCACTCTCTCCCGAGAAATACGAGGATGAAACCAATGACAGTT 614
Qy 541 TACAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGACCAACCAATGGCAGAGTCCA 600
Db 615 TACCAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGACCAACCAATGGCAGAGTCCA 674
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGACTCCGAGTTCGAATGAAAAT 660
Db 675 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGACTCCGAGTTCGAATGAAAAT 734
Qy 661 GAGCCACACTTCCATCATGTGTTAGCCAGAGAAAGAGGGGTCAAGGACTCTTCTCCCTGTGC 720
Db 735 GAGCCACACTTCCATCATGTGTTAGCCAGAGAAAGAGGGGTCAAGGACTCTTCTCCCTGTGC 794
Qy 721 GTAGCTTTTCATCTGTAAAGTTTCACTTTAAAGAGAGAAACTTCTTGGTCTCAAAACAAG 780
Db 795 GTAGCTTTTCATCTGTAAAGTTTCACTTTAAAGAGAGAAACTTCTTGGTCTCAAAACAAG 854
Qy 781 GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 840
Db 855 GAGATGGGCTCCCAAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCTCAAGAC 914
Qy 841 GGGAAAGCATCACTCATGAAGAAAGAGAGATTTTGGCTGAAGAGCTGTGTAATCTCTCCA 900
Db 915 GGGAAAGCATCACTCATGAAGAAAGAGAGATTTTGGCTGAAGAGCTGTGTAATCTCTCCA 974
Qy 901 GATCTGTGTGCTTTTGTGTGTGTAGAAATCTCAGATGAAGCTTTCATTCACCCATC 960
Db 975 GATCTGTGTGCTTTTGTGTGTGTAGAAATCTCAGATGAAGCTTTCATTCACCCATC 1034
Qy 961 TGTGAGAAATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCTGGCTTGGTG 1020
Db 1035 TGTGAGAAATGCCACCTTTTCAAGAGTACCAAGGAAAGGAGATGCCCTGGCTTGGTG 1094
Qy 1021 GTTCATGTGGCCCCAGCATCTGTGTTGTGGACAGAGTACCAGAGTGGATGAGAGG 1080
Db 1095 GTTCATGTGGCCCCAGCATCTGTGTTGTGGACAGAGTACCAGAGTGGATGAGAGG 1154
Qy 1081 TTTGGGCTGACACCCAGCATTTGTCTCTGAATGAGAACTGTGCCTCAGTTTCAACACTT 1140
Db 1155 TTTGGGCTGACACCCAGCATTTGTCTCTGAATGAGAACTGTGCCTCAGTTTCAACACTT 1214
Qy 1141 GCGAGCCACAAGATTTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGGTCT 1200
Db 1215 GCGAGCCACAAGATTTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGGTCT 1274
Qy 1201 ACCAGTTTCCGCTGAAGAGAGGGCCCCACCTCAGTGTGCCATGGTTCAGGGTGA 1260
Db 1275 ACCAGTTTCCGCTGAAGAGAGGGCCCCACCTCAGTGTGCCATGGTTCAGGGTGA 1334
Qy 1261 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1335 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1394
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGGCGTGCAGCTTCCCAATTCCTCAGCAGAGCGTG 1380
|||||

Db 1395 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTG 1454
Qy 1381 CAGAGTACAGGAGAGTGCAGGACGGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Db 1455 CAGAGTACAGGAGAGTGCAGGACGGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1514
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCACT 1500
Db 1515 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCACT 1574
Qy 1501 GCCACACTTGTCAAATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGAAGGACA 1560
Db 1575 GCCACACTTGTCAAATAAGCCCCGACACAGTCTCTGCTACTTGGACTGTGGTGAAGGACA 1634
Qy 1561 TTTGGGACAGTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTTGGGACACCTTGGCT 1620
Db 1635 TTTGGGACAGTGTGCCGTCAATTACGGAGACAGGTGACAGGGTCTTGGGACACCTTGGCT 1694
Qy 1621 GCTGTGTTTGTGTCCCACTGCGACGAGATCACACACGGGCTTGCCAAGTATCTTGTCTG 1680
Db 1695 GCTGTGTTTGTGTCCCACTGCGACGAGATCACACACGGGCTTGCCAAGTATCTTGTCTG 1754
Qy 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTCTGTGTGTGCTGCC 1740
Db 1755 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTCTGTGTGTGCTGCC 1814
Qy 1741 CCCAACAGCTCAAAGCTTGGCTCCAGAGTACCAACAACAGTCCAGGAGGCTCTGCAC 1800
Db 1815 CCCAACAGCTCAAAGCTTGGCTCCAGAGTACCAACAACAGTCCAGGAGGCTCTGCAC 1874
Qy 1801 CACATCAGTATGATTCTTGCCTCAAAATGCTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1860
Db 1875 CACATCAGTATGATTCTTGCCTCAAAATGCTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1934
Qy 1861 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTTGAAGAGTTCAGACCTGT 1920
Db 1935 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTTGAAGAGTTCAGACCTGT 1994
Qy 1921 CTGTGGGCACTGCAAGCATGCTTTGGCTTGGCTGCTGTGTCACACTCTGGCTGGAA 1980
Db 1995 CTGTGGGCACTGCAAGCATGCTTTGGCTTGGCTGCTGTGTCACACTCTGGCTGGAA 2054
Qy 1981 GTGTCTTATTTCCGGGACACCATGCCCTGCGAGCTCTGTGTCGGATGGGAAAGATGCC 2040
Db 2055 GTGTCTTATTTCCGGGACACCATGCCCTGCGAGCTCTGTGTCGGATGGGAAAGATGCC 2114
Qy 2041 ACCCTCTGTATACATGAAGCCACCTTGGAAAGATGTTTGAAGAGGAGACAGTGGAAAG 2100
Db 2115 ACCCTCTGTATACATGAAGCCACCTTGGAAAGATGTTTGAAGAGGAGACAGTGGAAAG 2174
Qy 2101 ACACAGCACAACTGCCAAGCCATCAGCGTGGGATGCGGATGAACGCGAGTTCAAT 2160
Db 2175 ACACAGCACAACTGCCAAGCCATCAGCGTGGGATGCGGATGAACGCGAGTTCAAT 2234
Qy 2161 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCAACTTCAGC 2220
Db 2235 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCAACTTCAGC 2294
Qy 2221 GAGAAAGTGGAGTGTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Db 2295 GAGAAAGTGGAGTGTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2354
Qy 2281 CCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGCAGATCGAGGAGATGGAGAG 2340
Db 2355 CCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGCAGATCGAGGAGATGGAGAG 2414
Qy 2341 GCGAGGAGAAAGCGGAGCTCGCGAGTGGGGCGGCTCTCTGTCCAGGAGAGTGGCA 2400
Db 2415 GCGAGGAGAAAGCGGAGCTCGCGAGTGGGGCGGCTCTCTGTCCAGGAGAGTGGCA 2474
Qy 2401 GCGGCTTGGAGATGGGGAGCTTCAGAGAAAGCGGGCCCCACAGAGGAGCCACAGGCC 2460
Db 2475 GCGGCTTGGAGATGGGGAGCTTCAGAGAAAGCGGGCCCCACAGAGGAGCCACAGGCC 2534
|||||

QY 2461 AAGAAGGTCAGAGCCAGTGA 2481
 |||||
 Db 2535 AAGAAGGTCAGAGCCAGTGA 2555

RESULT 3

BC001939

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC001939 2997 bp mRNA linear PRI 12-JUL-2001
 Homo sapiens, putative prostate cancer susceptibility protein,
 clone MGC:4102 IMAGE:2820640, mRNA, complete cds.
 BC001939
 BC001939.1 GI:12804972
 MGC.
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2997)
 Strausberg, R.
 Direct Submission
 Submitted (29-JAN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov

Tissue Procurement: DCTD/Drp

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Varon Butterfield,

Suanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 9 Row: i Column: 17

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 7022621.

FEATURES

source

Location/Qualifiers

1..2997

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:4102 IMAGE:2820640"

/tissue_type="Lung, small cell carcinoma"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

61..2541

/codon_start=1

/product="putative prostate cancer susceptibility protein"

/protein_id="AAH01939.1"

/db_xref="GI:12804973"

/translation="MWALCSLLRSAGRTMSGRTISOAPARRPRKDPRLHLRTRE

KRPSGSGGNTVYLVVAGSRDSGALVYFSEFNLYLNCGEGVRLQHEHLKV

ARLDNIFLRHNSVNGSLMILTKETGLPKVLSGPPQLEKYLEAKIFPSGLKG

IELAVRHSAPEYEDMTVQIPIHSEQRKHQWSPERPLSRSPERSSELN

ENEPHLPHGSQRGRVDSLVVAFICKLHKGKFLVLRKAKEMGLPVGTAAIPIIA

AVKDGKSLTHEGRIELABELCTPPDGAFFVVECPDESFIQICENATFORYOGKAD

APVALVHMAPASVLVDVRYQOMERFQDTHLVLENCAVSNLNRSHKIQTLNLI

HPDIFPLITSFCKEGLTSLVPMVQEGCLLYQLRPRREWRDAIITCNFEFIVEA

LQLPNQSVQYRRSAQDGPAERKQYPIEIIPLGTGSAIPMKIRNVSATLVNISP

DTSLLLDCGEGTGGOLCRHYGDQVDRVLGTAAVFEVSHLHADHTGLPSSILLQERAL

BASE COUNT 740 a 806 c 853 g 598 t
 ORIGIN

Query Match 99.9%; Score 2477.8; DB 9; Length 2997;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGGGCGGAGCAGCACCATTGTCGAGGAGCG 60

Db 61 ATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGGGCGGAGCAGCACCATTGTCGAGGAGCG 120

QY 61 ACCATATCGCAGGACCCGCGCGCGGAGCGGCGCGGAGCAGCACCATTGTCGAGGAGCG 120

Db 121 ACCATATCGCAGGACCCGCGCGCGGAGCGGCGCGGAGCAGCACCATTGTCGAGGAGCG 180

QY 121 CGACGCGAGAGAACCGCGGACCCGCTCGGCGGCGGAGCAGCACCATTGTCGAGGAGCG 180

Db 181 CGACGCGAGAGAACCGCGGACCCGCTCGGCGGCGGAGCAGCACCATTGTCGAGGAGCG 240

QY 181 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGGCGGAGCAGCACCATTGTCGAGGAGCG 240

Db 241 CAGGTGGTGGCAGCGGCTAGCGGAGCTCGGCGGCGGAGCAGCACCATTGTCGAGGAGCG 300

QY 241 AACCGGTATCTTCAACTGTGGAGAGCGTTTTCAGAGACTCATCGAGGAGCACAAGTTA 300

Db 301 AACCGGTATCTTCAACTGTGGAGAGCGTTTTCAGAGACTCATCGAGGAGCACAAGTTA 360

QY 301 AAGGTTGCTCGCTGGACCAACATATTTCTGACACCAATATCTTCTGCTTCTTCTGCGGCG 360

Db 361 AAGGTTGCTCGCTGGACCAACATATTTCTGACACCAATATCTTCTGCTTCTTCTGCGGCG 420

QY 361 TTAAGTGAATGATTCTTACTTTAAAGGAACCGGGCTTCAAAGTGTGACTTCTTCTGGA 420

Db 421 TTAAGTGAATGATTCTTACTTTAAAGGAACCGGGCTTCAAAGTGTGACTTCTTCTGGA 480

QY 421 COTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAAATTTTCTGCTCCATTGAAAGGA 480

Db 481 COTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAAATTTTCTGCTCCATTGAAAGGA 540

QY 481 ATAGAACTGGGTGTGCGGCGCCCACTCTGCCCGAGAAATACGAGGATGAACCAATGACAGTT 540

Db 541 ATAGAACTGGGTGTGCGGCGCCCACTCTGCCCGAGAAATACGAGGATGAACCAATGACAGTT 600

QY 541 TACCAGATCCCCATACACAGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 600

Db 601 TACCAGATCCCCATACACAGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 660

QY 601 GAAAGGCTCTTCAGAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAGTGAATGAAT 660

Db 661 GAAAGGCTCTTCAGAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAGTGAATGAAT 720

QY 661 GAGCCACACCTTCCACATAGTGTGTAGCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGT 720

Db 721 GAGCCACACCTTCCACATAGTGTGTAGCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGT 780

QY 721 GTAGTCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAATCTTCTGCTGCTCAAGCAAG 780

Db 781 GTAGTCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAATCTTCTGCTGCTCAAGCAAG 840

QY 781 GAGATGGGCTCCCAAGTGGGAGCTGCCATCGCTCCCATCTCCATCTTCTGCTCTCAGGAG 840

Db 841 GAGATGGGCTCCCAAGTGGGAGCTGCCATCGCTCCCATCTCCATCTTCTGCTCTCAGGAG 900

QY 841 GGGAAAAAGCATCACTCATGAAGAGAGAGATTTTGGGTGAAGAGAGTGTGTACTCTCTCCA 900

Db 901 GGGAAAAAGCATCACTCATGAAGAGAGAGATTTTGGGTGAAGAGAGTGTGTACTCTCTCCA 960

QY 901 GATCCTGGTGTGCTTTTGTGGTAGAATGTCCAGATGAAAGCTTCTTCAACCCATC 960

Db 961 GATCTCTGCTGCTTTTGTGTGTGTAGATGTCCAGATGAAGCTTATTCAACCCATC 1020
Qy 961 TGTGAAATGACACCTTTTCAGAGGTACCAAGAAAGGAGATGCCCGCGCTGTGTG 1020
Db 1021 TGTGAAATGACACCTTTTCAGAGGTACCAAGAAAGGAGATGCCCGCGCTGTGTG 1080
Qy 1021 GTTTCATGTGGCCCGCAGCATCTGTGCTGTGGACAGAGGTACCCAGCAGTGTGAGAGG 1080
Db 1081 GTTTCATGTGGCCCGCAGCATCTGTGCTGTGGACAGAGGTACCCAGCAGTGTGAGAGG 1140
Qy 1081 TTTTGGCCCTGACACCCAGCAGTGTGCTGTGAATGAGAACTGTGCTCAGTTCACAACTT 1140
Db 1141 TTTTGGCCCTGACACCCAGCAGTGTGCTGTGAATGAGAACTGTGCTCAGTTCACAACTT 1200
Qy 1141 CGAGCCCAAGATTCAAAACCCAGCTCAACCTTCATCCACCGGACATTTTCCCCCTGTC 1200
Db 1201 CGAGCCCAAGATTCAAAACCCAGCTCAACCTTCATCCACCGGACATTTTCCCCCTGTC 1260
Qy 1201 ACCAGTTTCCGCTGTAAGNAGGAGGGCCCCACCTCTCAGTGTGCCATGGTTTCAGGGTGA 1260
Db 1261 ACCAGTTTCCGCTGTAAGNAGGAGGGCCCCACCTCTCAGTGTGCCATGGTTTCAGGGTGA 1320
Qy 1261 TGCCTCTCAAGTACCAAGTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db 1321 TGCCTCTCAAGTACCAAGTCCGCTCCAGAGGGAGTGGCAGAGGGATGCCATTATTACT 1380
Qy 1321 TGCAATCTGTAGGAATTCATAGTTGAGGCGTGCAGCTTCCAACTTCAGCAGAGCGTG 1380
Db 1381 TGCAATCTGTAGGAATTCATAGTTGAGGCGTGCAGCTTCCAACTTCAGCAGAGCGTG 1440
Qy 1381 CAGGAGTACAGGAGAGTGCAGCAGAGCGGCCAGCCCGAGCCAGCAGAGAGAAAGTCACTAC 1440
Db 1441 CAGGAGTACAGGAGAGTGCAGCAGAGCGGCCAGCCCGAGCCAGCAGAGAGAAAGTCACTAC 1500
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGCTGTGCATCCCGATCCCGATTCGAAATCTCAGT 1500
Db 1501 CCAGAAATCATCTTCTTGGAAACAGGCTGTGCATCCCGATTCGAAATCTCAGT 1560
Qy 1501 GCCACATCTGTCAACATTAAGCCCGACACGCTCTCTGTACTTGGATCTGTGAGGGCACA 1560
Db 1561 GCCACATCTGTCAACATTAAGCCCGACACGCTCTCTGTACTTGGATCTGTGAGGGCAGC 1620
Qy 1561 TTTTGGCAGCTGTGCGCTATTACGAGACAGGTGGACAGGGTCTTGGCCACCTGGCT 1620
Db 1621 TTTTGGCAGCTGTGCGCTATTACGAGACAGGTGGACAGGGTCTTGGCCACCTGGCT 1680
Qy 1621 GCTGTGTTTGTGTCCTCCACTGCAACGAGATCACACAGGGCTTCCCAAGTATCTTGCTG 1680
Db 1681 GCTGTGTTTGTGTCCTCCACTGCAACGAGATCACACAGGGCTTCCCAAGTATCTTGCTG 1740
Qy 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGGAAAGCGCTTACCCCTTGTGTTGTTGCC 1740
Db 1741 CAGAGAGAACCGCCTTGGCATCTTTTGGGAAAGCGCTTACCCCTTGTGTTGTTGCC 1800
Qy 1741 CCCAACGAGCTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1800
Db 1801 CCCAACGAGCTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGAGGCTCTGCAC 1860
Qy 1801 CACATCAGTATGATTCCTGCCAAATTCCTTCCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1861 CACATCAGTATGATTCCTGCCAAATTCCTTCCAGGAGGGCTGAGATCTCCAGTCTGCA 1920
Qy 1861 GTGGAAGATTGATCAGTTGCTGTGTGGCAACATGTGATTTGGAAGAGTTTCAGACCTGT 1920
Db 1921 GTGGAAGATTGATCAGTTGCTGTGTGGCAACATGTGATTTGGAAGAGTTTCAGACCTGT 1980
Qy 1921 CTGGTGGCCACTGCAAGCATCGTTTGGCTGTGGCTGGTGTGCAACCTCTGGCTGGAA 1980
Db 1981 CTGGTGGCCACTGCAAGCATCGTTTGGCTGTGGCTGGTGTGCAACCTCTGGCTGGAA 2040
Qy 1981 GTGGTCTATTCCGGGGACACCATGCCCTCGCAGGCTCTGGTCCGGATGGGAAAGATGCC 2040

Db 2041 GTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2100
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGNAGGAGCAGTGGAAAG 2100
Db 2101 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGNAGGAGCAGTGGAAAG 2160
Qy 2101 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2160
Db 2161 ACACACAGCAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2220
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCGCTTTCAGCCCACTTCAAG 2220
Db 2221 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCGCTTTCAGCCCACTTCAAG 2280
Qy 2221 GAGAAAGTGGAGTTGCTTTTGACCATGAAGTCTGCTTTGAGAGCTTTTCCAACTATG 2280
Db 2281 GAGAAAGTGGAGTTGCTTTTGACCATGAAGTCTGCTTTGAGAGCTTTTCCAACTATG 2340
Qy 2281 CCCAAGCTGATTTCCCGCACTGAAAGCCCTGTTTCTGCGCACATCGAGGAGATGGAGG 2340
Db 2341 CCCAAGCTGATTTCCCGCACTGAAAGCCCTGTTTCTGCGCACATCGAGGAGATGGAGG 2400
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGCAGGTGCGGCGCGCCCTCTCTGTCCAGGAGCTGGCA 2400
Db 2401 CGCAGGAGAAAGCGGAGCTGCGCAGGTGCGGCGCGCCCTCTCTGTCCAGGAGCTGGCA 2460
Qy 2401 GCGCGCTGAGGATGGGAGCCCTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Db 2461 GCGCGCTGAGGATGGGAGCCCTCAGCAGAAAGCGGCGCCACACAGAGGAGCCACAGGCC 2520
Qy 2461 AAGAAGTCTCAGAGCCCACTGTA 2481
Db 2521 AAGAAGTCTCAGAGCCCACTGTA 2541

BC004158 3006 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, putative prostate cancer susceptibility protein,
clone MGC:2441 IMAGE:2820640, mRNA, complete cds.
BC004158
ACCESSION BC004158
VERSION BC004158.1 GI:13278770
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3006)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kattman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 2 Row: C Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7022621.
Location/Qualifiers
1. .3006

Db 121 CCGACGCGAGAGAGCGGACCGTGGGGTGCTCCGGCGGCCCAACACCGGTGTACCTG 180
Qy 181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGCGCGCGGCTCTAGCTCTTCTCCAGTTTC 240
Db 181 CAGGTGGTGGCAGCGGGTAGCCGGGACTCGGCGCGCGGCTCTAGCTCTTCTCCAGTTTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGAGAGACACAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCAATTCAGAGACTCATGAGAGACACAAGTTA 300
Qy 301 AAGGTGGTGGTGGCAGCAACATATCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
Db 301 AAGGTGGTGGTGGCAGCAACATATCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
Qy 361 TTAAGTGGAAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTCGA 420
Db 361 TTAAGTGGAAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTCGA 420
Qy 421 CTTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATGTAAGGA 480
Db 421 CTTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATGTAAGGA 480
Qy 481 ATAGAACTGGCTGGCGGCCCACTCTGCCAGATACGAGATCAACCATGACAGTT 540
Db 481 ATAGAACTGGCTGGCGGCCCACTCTGCCAGATACGAGATCAACCATGACAGTT 540
Qy 541 TACCAAGTCCCCATACAGTGAACAGAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
Db 541 TACCAAGTCCCCATACAGTGAACAGAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAAAAT 660
Qy 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Qy 721 GTAGCTTTTCATCTGTAAAGTTCACCTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Db 721 GTAGCTTTTCATCTGTAAAGTTCACCTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCTGCTGTCAAGGAC 840
Db 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCTGCTGTCAAGGAC 840
Qy 841 GGGAAAGCATCACTCATGAAGGAGAGATTTGGCTGAAGAGCTGTACTCTCCA 900
Db 841 GGGAAAGCATCACTCATGAAGGAGAGATTTGGCTGAAGAGCTGTACTCTCCA 900
Qy 901 GATCTGTGTGTCTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
Db 901 GATCTGTGTGTCTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCAACCCATC 960
Qy 961 TGTGAAATGCCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGGTGGTG 1020
Db 961 TGTGAAATGCCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCGGTGGTG 1020
Qy 1021 GTTCATGGCCCGCAGCATCTGTCTGTGGACAGCAGGTACGAGCAGTGGATGGAGG 1080
Db 1021 GTTCATGGCCCGCAGAAATCTGTGTCTGTGGACAGCAGGTACGAGCAGTGGATGGAGG 1080
Qy 1081 TTTGGSCCTGACACCCAGCACTTGTCTGTGAATGAGAACTGTGCCTCAGTTCACACCTT 1140
Db 1081 TTTGGSCCTGACACCCAGCACTTGTCTGTGAATGAGAACTGTGCCTCAGTTCACACCTT 1140
Qy 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTCTC 1200
Db 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGTCTC 1200
Qy 1201 ACCAGTTTCGGTGTAAAGAGAGGGGCCACCCCTCAGTGTGCCCATGTTTTCAGGGTGA 1260

Db 1201 ACCAGTTTCGGTGTAAAGAGAGGGGCCACCCCTCAGTGTGCCCATGGTTTCAGGGTGA 1260
Qy 1261 TGCCTCTCTCAAGTACCAGTCCGTCCAGAGGAGGTGGCAGAGGATGCCATTATTTACT 1320
Db 1261 TGCCTCTCTCAAGTACCAGTCCGTCCAGAGGAGGTGGCAGAGGATGCCATTATTTACT 1320
Qy 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGCTGAGCTTCCCACTTCCAGCAGAGCGTG 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGCTGAGCTTCCCACTTCCAGCAGAGTG 1380
Qy 1381 CAGGAGTACAGAGAGGTGCGCAGGACGCGCCAGCCCGCAGCAGAGAGAAAAGTCAGTAC 1440
Db 1381 CAGGAGTACAGAGAGGTGCGCAGGACGCGCCAGCCCGCAGCAGAGAGAAAAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCA 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCA 1500
Qy 1501 GCCACACTTGTCAACATGAAGCCCGACAGCTCTCTGCTACTGGACTGTGGTGAGGGCAG 1560
Db 1501 GCCACACTTGTCAACATGAAGCCCGACAGCTCTCTGCTACTGGACTGTGGTGAGGGCAG 1560
Qy 1561 TTTGGGCACTGTGCGCTCATTACGAGACACAGGTGGACAGGGTCTCTGGGCACTCGGCT 1620
Db 1561 TTTGGGCACTGTGCGCTCATTACGAGACACAGGTGGACAGGGTCTCTGGGCACTCGGCT 1620
Qy 1621 GCTGTGTTTGTCTCCACCTGCACGAGATCACACACGGGCTTCCCAAGTATCTTGCTG 1680
Db 1621 GCTGTGTTTGTCTCCACCTGCACGAGATCACACACGGGCTTCCCAAGTATCTTGCTG 1680
Qy 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCCCTTTCACCTTTGCTGGTGGTGGC 1740
Db 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCCCTTTCACCTTTGCTGGTGGTGGC 1740
Qy 1741 CCCAACACAGCTCAAGCTGCTCCAGCAGTACCACACAGTGCAGAGGTCTCTGCAC 1800
Db 1741 CCCAACACAGCTCAAGCTGCTCCAGCAGTACCACACAGTGCAGAGGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTTCTGCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTTGA 1860
Db 1801 CACATCAGTATGATTTCTGCCAAATGCCTTCAGGAAGGGGTGAGATCTCCAGTCTTGA 1860
Qy 1861 GTGGAAGATGATCAGTTGCTGTGGCAATGTGATTTGGAAAGAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATGATCAGTTGCTGTGGCAATGTGATTTGGAAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGGTGGCGCATGCAAGCATGCTGCTGGCTGGTGCACACCTCTCGCTGGAAA 1980
Db 1921 CTGGTGGCGCATGCAAGCATGCTGCTGGCTGGTGCACACCTCTCGCTGGAAA 1980
Qy 1981 GTGGTCTATTCCGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGTGC 2040
Db 1981 GTGGTCTATTCCGGGACACCATGCCCTCGAGGCTCTGGTCCGGATGGGAAAGTGC 2040
Qy 2041 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGCAGTGGAAAAG 2100
Db 2041 ACCCTCTGTATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGCAGTGGAAAAG 2100
Qy 2101 ACACACAGCAACGTCACAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2160
Db 2101 ACACACAGCAACGTCACAGCCATCAGCGTGGGATGCGGATCAACCGGAGTTCATT 2160
Qy 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Qy 2221 GAGAAAGTGGAGTTGCTTTGAACCATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2280
Db 2221 GAGAAAGTGGAGTTGCTTTGAACCATGAAGTCTGCTTTGGAGACTTTTCCAACTATG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340

```

Qy 2341 CGCAGGAGAGCGGAGCTCGGCAGCTGCGGCGGCGCCCTCTGTCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGCGGAGCTCGGCAGCTGCGGCGGCGCCCTCTGTCAGGAGCTGGCA 2400
Qy 2401 GCGCGCTTGGAGGATGGGAGCTCAGCAGAGAGCGGCGCCACACAGAGAGGAGCCAGAGGCC 2460
Db 2401 GCGCGCTTGGAGGATGGGAGCTCAGCAGAGAGCGGCGCCACACAGAGAGGAGCCAGAGGCC 2460
Qy 2461 AAGAAGGTCAGAGCCAGTGA 2481
Db 2461 AAGAAGGTCAGAGCCAGTGA 2481

RESULT 6
AK001392 Homo sapiens cDNA FLJ10530 fis, clone NT2RP2000985. linear PRI 01-AUG-2002
LOCUS AK001392 2976 bp mRNA
DEFINITION Homo sapiens cDNA FLJ10530 fis, clone NT2RP2000985.
ACCESSION AK001392
VERSION AK001392.1 GI:7022621
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,
clone lib:NT2RP2 clone:NT2RP2000985.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuo,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2976)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1. .2976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2000985"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/feature="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
75. .2639
/feature="protein product"
/codon_start=1
/protein_id="BAA91666.1"
/db_xref="GI:7022622"
/translation="MWALCSLLRGAAGRTMSOGRITISQAPARRPRPKDLRLHRTRE
KRPGSGGNTVYLVQVAGSRDGAALVFSEFNRYLPNCEGVQRLMOEHKLAV
ARLDNIPLTRMHNWNGLSMILTKEGLPKVLSGPQLEKYLAIKIFSPKLG
TELARPHSAPEDEMTVYQIPHSEQRKHQWPQSPERLURLSPKSSDSEN
ENEPHLPHGVSRQGRVDSLSLVAFIKLHLKRNFLVLKAKENGLPVGTAAAPIIA
AVKDKGKITHREGRIELAEELCTPPDGAAFVVECPDESEFTQICENATFORVQKAD
APVALVVMAPASVLDSRYQWMERFGPDQHLVLNENCASVHNLRSKHTQIOLNI
HPDIFPLLTSPRCKEGPTLSVPVQGECLLYQLRPREWQDAILITCNPEEFVEAL

```

```

OLPNFOOSVQVEYRRSAQDGPAPAEKRSQYPEIIFLGTGSAIPMKIRNVSATLVNISPD
TSLLLCGCGTFCGLCRHYGQDOVRLVLTAAVFSVHLHADHTGLPSILLORERALA
SLGKPLRPLLVAPNOLKAWLOQYVNCQVLEHHISMIPAKCIQEGNEISSPAVERLI
SSLURTCDLLEFPOTCLVRICKHAFGCALVHSGKVYSGDTWPCFALVRMGKDATLL
IHEATLDEGLBEEAVEKTHSTTSQAI SVGRMNAEFIMLNHFQRYAKVFLFSPNSE
KVGVAFDHMKVCLGDFTPMKLIPPLKALFAGDI EEMEERERKRELRQVRAALLSREL
AGGLEDSLSRSQPTQSRPRSEPSSEDLGPELRRLCLVFCPTHAPVSALLAGRS"
BASE COUNT 712 a 807 c 856 g 501 t
ORIGIN
Query Match 98.7%; Score 2448.2; DB 9; Length 2976;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 3; Indels 4; Gaps 2;
Qy 1 ATGTGGGGGCTTTGCTCGCTGCTGCGGTCCGCGCGCGGACGACCATGTGCGCAGGAGCC 60
Db 75 ATGTGGGGGCTTTGCTCGCTGCTGCGGTCCGCGCGCGGACGACCATGTGCGCAGGAGCC 134
Qy 61 ACCATATCGCAGGACCCGCGCGCGGAGCGCGCGGCAAGGACCCGCTCGCGCACCTG 120
Db 135 ACCATATCGCAGGACCCGCGCGCGGAGCGCGCGGCAAGGACCCGCTCGCGCACCTG 194
Qy 121 CGCACGCGAGAGAACGCGGACCGTCCGGGGTGTCTCCGCGCGCCCAACACCGTGTACCTG 180
Db 195 CGCACGCGAGAGAACGCGGACCGTCCGGGGTGTCTCCGCGCGCCCAACACCGTGTACCTG 254
Qy 181 CAGTGTGTGCGAGGGGTAGCGGGACTCGGCGCGCGGCTCTAGCTTCTTCGAGTTTC 240
Db 255 CAGTGTGTGCGAGGGGTAGCGGGACTCGGCGCGCGGCTCTAGCTTCTTCGAGTTTC 314
Qy 241 AACCGGTATCTTCTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGACACAAAGTTA 300
Db 315 AACCGGTATCTTCTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGACACAAAGTTA 374
Qy 301 AAGTTTGTGCGCTCGCAACATATTCCTGCACAGAAATGCACTCGTCTTAATGTGGGGGC 360
Db 375 AAGTTTGTGCGCTCGCAACATATTCCTGCACAGAAATGCACTCGTCTTAATGTGGGGGC 434
Qy 361 TTAAGTGAATGATCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTCTGGA 420
Db 435 TTAAGTGAATGATCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTCTGGA 494
Qy 421 CCTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTTGAAAGGA 480
Db 495 CCTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTTGAAAGGA 554
Qy 481 ATAGAACTGGTGTGCGGCGCCCTCTCTGCCCCAGAAATACGAGGATGAAACCATGACAGTT 540
Db 555 ATAGAACTGGTGTGCGGCGCCCTCTCTGCCCCAGAAATACGAGGATGAAACCATGACAGTT 614
Qy 541 TACCAGATCCCATACAGTGAACAGAGAGGGGAAAGCAACCAATGCGAGAGTCCA 600
Db 615 TACCAGATCCCATACAGTGAACAGAGAGGGGAAAGCAACCAATGCGAGAGTCCA 674
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db 675 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 734
Qy 661 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGAGGGGTGAGGAGCTTCTCCCTGGTC 720
Db 735 GAGCCACACCTTCCACATGGTGTGTAGCCAGAGAGAGGGGTGAGGAGCTTCTCCCTGGTC 794
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCTCAAGCAAG 780
Db 795 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAGAACTTCTTGGTGTCTCAAGCAAG 854
Qy 781 GAGATGGGCTCCAGTGTGGGACAGCTGCCATCTCCCATCATTTGCTGTCTCAAGGAC 840
Db 855 GAGATGGGCTCCAGTGTGGGACAGCTGCCATCTCCCATCATTTGCTGTCTCAAGGAC 914
Qy 841 GGGAAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCA 900
Db 915 GGGAAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCA 974

```

QY 901 GATCTCTGTCGCTCTTTTGTGTGTGTAGTAATGTCCAGATGAAGCTTCAATCAACCCATC 960
DB 975 GATCTCTGTCGCTCTTTTGTGTGTGTAGTAATGTCCAGATGAAGCTTCAATCAACCCATC 1034
QY 961 TGTGAGAAATGCCACCTTTCCAGAGGTACCAAGAAAGGAGAGATGCCCGTGTGCTTGGTG 1020
DB 1035 TGTGAGAAATGCCACCTTTCCAGAGGTACCAAGAAAGGAGAGATGCCCGTGTGCTTGGTG 1094
QY 1021 GTTCATATGGCCCCCAGCACTCTGTCTGTGTGTGACAGCAGGTACCAAGGTGATGAGAGG 1080
DB 1095 GTTCATATGGCCCCCAGCACTCTGTCTGTGTGTGACAGCAGGTACCAAGGTGATGAGAGG 1154
QY 1081 TTTGGGCTTGACACCCAGCACTTGTCTCTGAATGAGAACTGTGCTCAGTTCACAACTTT 1140
DB 1155 TTTGGGCTTGACACCCAGCACTTGTCTCTGAATGAGAACTGTGCTCAGTTCACAACTTT 1214
QY 1141 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTGCTC 1200
DB 1215 CGCAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTGCTC 1274
QY 1201 ACCAGTTTCCGCTGTAAAGAGGAGGCCCCACCTCAGTGTGCCATGTGTTCCAGGTTGAA 1260
DB 1275 ACCAGTTTCCGCTGT---AAGGAGGCCCCACCTCAGTGTGCCATGTGTTCCAGGTTGAA 1331
QY 1261 TGCCTCTCAAGTACCAAGTCCGCTCCAGGAGGAGTGGCAGAGGATGCCATTTATTACT 1320
DB 1332 TGCCTCTCAAGTACCAAGTCCGCTCCAGGAGGAGTGGCAGAGGATGCCATTTATTACT 1391
QY 1321 TGCATCTCTGAGAAATTCATAGTTGAGGCGCTGCGAGCTTCCCACTTCCAGCAGAGCGTG 1380
DB 1392 TGCATCTCTGAGAAATTCATAGTTGAGGCGCTGCGAGCTTCCCACTTCCAGCAGAGCGTG 1451
QY 1381 CAGGAGTACAGGAGGAGTGGCAGGACCGCCAGCGCCAGCCAGCAGAGAAAGAGTCAGTAC 1440
DB 1452 CAGGAGTACAGGAGGAGTGGCAGGACCGCCAGCGCCAGCCAGCAGAGAAAGAGTCAGTAC 1511
QY 1441 CCAGAAATCATCTCTTGGAAACAGGCTCTGCCATCCCATCCGATGAAGATTGCAATGTCAGT 1500
DB 1512 CCAGAAATCATCTCTTGGAAACAGGCTCTGCCATCCCATCCGATGAAGATTGCAATGTCAGT 1571
QY 1501 GCCACACTTGTCAACATAGCCCGACAGCTCTCTGCTACTGACTGTGTCAGGAGGACCA 1560
DB 1572 GCCACACTTGTCAACATAGCCCGACAGCTCTCTGCTACTGACTGTGTCAGGAGGACCA 1631
QY 1561 TTTGGGCACTGTGCGCTCATTAACGAGACCGAGTGGCAGGCTCTGGGCACTTGGCT 1620
DB 1632 TTTGGGCACTGTGCGCTCATTAACGAGACCGAGTGGCAGGCTCTGGGCACTTGGCT 1691
QY 1621 GTGTGTTTGTGTCCACTGTCAGCAGATCACACACCGGCTTGCCTTGTGCTGCTG 1680
DB 1692 GTGTGTTTGTGTCCACTGTCAGCAGATCACACACCGGCTTGCCTTGTGCTGCTG 1751
QY 1681 CAGAGAGAGCGGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGCTGCTG 1740
DB 1752 CAGAGAGAGCGGCTTGGCATCTTTGGAAAGCGCTTACCCCTTGTGCTGCTGCTG 1811
QY 1741 CCCAACCACTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGGAGTCTCTGCAC 1800
DB 1812 CCCAACCACTCAAGCCTGGCTCCAGCAGTACTACCAACAGTGCAGGAGTCTCTGCAC 1871
QY 1801 CACATCAGTATGATTCCTGCCAAATGCCCTTCCAGGAGGGGCTGAGATCTCCAGTCTCGCA 1860
DB 1872 CACATCAGTATGATTCCTGCCAAATGCCCTTCCAGGAGGGGCTGAGATCTCCAGTCTCGCA 1931
QY 1861 GTGGAAGATTGATCAGTTCGCTGTGTCGCAACATGATTTGGAAGAGTTTCAGACCTGT 1920
DB 1932 GTGGAAGATTGATCAGTTCGCTGTGTCGCAACATGATTTGGAAGAGTTTCAGACCTGT 1991
QY 1921 CTGGTCGGCACTGCAAGCATGGTTTGGCTGTGCGCTGTCAGACCTCTCGCTGGA 1980
DB 1992 CTGGTCGGCACTGCAAGCATGGTTTGGCTGTGCGCTGTCAGACCTCTCGCTGGA 2051

QY 1981 GTGCTTATTTCCGGGACACCATGCCCTTCCGAGGCTCTGCTCCGATGGGAAAGATGCC 2040
DB 2052 GTGCTTATTTCCGGGACACCATGCCCTTCCGAGGCTCTGCTCCGATGGGAAAGATGCC 2111
QY 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAAAGCAGTGGAAAG 2100
DB 2112 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAAAGCAGTGGAAAG 2171
QY 2101 ACACACAGCAAAAGCTCCCAAGCCATCAGCTGGGATGCGGATGAACGCGAGTTCAAT 2160
DB 2172 ACACACAGCAAAAGCTCCCAAGCCATCAGCTGGGATGCGGATGAACGCGAGTTCAAT 2231
QY 2161 ATGCTGAACCACTTCAGCAGGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2220
DB 2232 ATGCTGAACCACTTCAGCAGGCTATGCCAAGTCCCTCTTCAGCCCACTTCAGC 2291
QY 2221 GAGAAAGTGGGAGTTGCTTTGACACATGAAGTCTCTTTGGAGATCTTTCCAAACAATG 2280
DB 2292 GAGAAAGTGGGAGTTGCTTTGACACATGAAGTCTCTTTGGAGATCTTTCCAAACAATG 2351
QY 2281 CCCAAGCTGATTTCCCTCACTGAAAGCCCTGTTGCTGCGCATCGAGGAGATGAGGAG 2340
DB 2352 CCCAAGCTGATTTCCCTCACTGAAAGCCCTGTTGCTGCGCATCGAGGAGATGAGGAG 2411
QY 2341 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGCGGCTCTCTTCAGGAGCTGSCA 2400
DB 2412 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGCGGCTCTCTTCAGGAGCTGSCA 2471
QY 2401 GCGGCTTGGAGAGTGGGAGCTCAGCAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2460
DB 2472 GCGGCTTGGAGAT-GGAGGCTCAGCAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2530
QY 2461 AAGAGGTTCAGAGCCCACTGA 2481
DB 2531 AAGAGGTTCAGAGCCCACTGA 2551

RESULT 7
AF308694
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AF308694
Gorilla gorilla ELAC2 (ELAC2) mRNA, complete cds.
AF308694.1 GI:10946488
Gorilla gorilla.
Gorilla gorilla.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
1 (bases 1 to 2893)
Tavtigian, S.V., Simard, J., Teng, D.H.F., Abtin, V., Baumgard, M.,
Beck, A., Camp, N.J., Carillo, A.R., Chen, Y., Dayananth, P.,
Desrochers, M., Dumont, M., Farnham, J.M., Frank, D., Frye, C.,
Gaffari, S., Gupta, J.S., Hu, R., Iliev, D., Janecki, T., Kort, E.N.,
Laity, K.E., Leavitt, A., Leblanc, G., McArthur-Morrison, J.,
Pederson, A., Penn, B., Peterson, K.T., Reid, J.E., Richards, S.,
Schroeder, M., Smith, R., Snyder, S.C., Swedlund, B., Swensen, J.,
Thomas, A., Tranchant, M., Woodland, A.M., Labrie, F., Skolnick, M.H.,
Neuhausen, S., Rommens, J., and Cannon-Albright, L.A.
A candidate prostate cancer susceptibility gene at chromosome 17p
Nat. Genet. 27 (2), 172-180 (2001)
21096977
2 (bases 1 to 2893)
Camp, N.J., Simard, J., Teng, D.H.F., Baumgard, M., Beck, A.,
Tavtigian, S.V., Chen, Y., Dayananth, P., Desrochers, M.,
Dumont, M., Farnham, J.M., Frank, D., Frye, C., Ghaffari, S.,
Gupte, J.S., Hu, R., Iliev, D., Janecki, T., Kort, E.N., Laity, K.E.,
Leavitt, A., Leblanc, G., McArthur-Morrison, J., Pederson, A., Penn, B.,
Peterson, K.T., Reid, J.E., Richards, S., Schroeder, M., Smith, R.,
Snyder, S.C., Swedlund, B., Swensen, J., Thomas, A., Tranchant, M.,
Woodland, A.M., Labrie, F., Skolnick, M.H., Neuhausen, S., Rommens, J.,
and Cannon-Albright, L.A.
Gorilla gorilla ortholog of human HPC2/ELAC2


```
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGCTGCTCCATCCCATGAATTCGAAATGTCAGT 1500
Qy 1501 GCCACACTTGTCAACATAGCCCGACACAGCTCTCTGCTACTGCACTGTGGTGAAGGCACA 1560
Db 1501 GCCACACTTGTCAACATAGCCCGACACAGCTCTCTGCTACTGCACTGTGGTGAAGGCACG 1560
Qy 1561 TTTGGGAGCTGTGGCTCATTAACGAGACACAGGTGAGCAGGCTCTGGGCAACCTGTGCT 1620
Db 1561 TTTGGGAGCTGTGGCTCATTAACGAGACACAGGTGAGCAGGCTCTGGGCAACCTGTGCT 1620
Qy 1621 GCTGTGTTTGTCTCCACTGCAACGAGATCACCACAGGCTTCCAGAGTATCTTGTG 1680
Db 1621 GCTGTGTTTGTCTCCACTGCAACGAGATCACCACAGGCTTCCAGAGTATCTTGTG 1680
Qy 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGAAAGCCGCTTACCCCTTTGCTGTGTGCT 1740
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGAAAGCCCTTACCCCTTTGCTGTGTGCT 1740
Qy 1741 CCCAACAGCTCAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Db 1741 CCCAACAGCTCAAGCTGTGCTCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCCTGCCAATGCCCTTACAGAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAATGCCCTTACAGAGGGCTGAGATCTCCAGTCTGCA 1860
Qy 1861 GTGGAAGATTCATCAGTTCGCTGTTCGCAACATGTATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTCATCAGTTCGCTGTTCGCAACATGTATTTGGAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGGTGCAGCTCAAGCATCGTTTGGCTGTGGCTGTGCACACCTCTCGCTGGAAA 1980
Db 1921 CTGGTGCAGCTCAAGCATCGTTTGGCTGTGGCTGTGCACACCTCTCGCTGGAAA 1980
Qy 1981 GTGGTCTATTCGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGGTCTATTCGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGAGCAAGCAGTGAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGAGCAAGCAGTGAAAG 2100
Qy 2101 ACACAGACACACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACCGGAGTTCAAT 2160
Db 2101 ACACAGACACACGTCCTCAAGCCATCAGCGTGGGATCGGATGAACCGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTACGCCCAACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTACGCCCAACTTCAC 2220
Qy 2221 GAGAAAGTGGGAGTTGCCCTTGACCAATGAAGTCTGCTTTGGAGACTTTTCAACAATG 2280
Db 2221 GAGAAAGTGGGAGTTGCCCTTGACCAATGAAGTCTGCTTTGGAGACTTTTCAACAATG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGGGACATCAGAGAGATGGAGGAG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGGGACATCAGAGAGATGGAGGAG 2340
Qy 2341 CGGAGGAGAGCGGAGCTGCGGAGGTGCGGCGGCTCTCTGTCCAGGAGCTGGCA 2400
Db 2341 CGGAGGAGAGCGGAGCTGCGGAGGTGCGGCGGCTCTCTGTCCAGGAGCTGGCA 2400
Qy 2401 GCGGCTCTGAGAGTGGGAGCTCAGCAGAGAGCGGCGCCACACAGGAGGCCACAGGCC 2460
Db 2401 GCGGCTCTGAGAGTGGGAGCTCAGCAGAGAGCGGCGCCACACAGGAGGCCACAGGCC 2460
Qy 2461 AAGAAGTTCAGAGCCCAAGTGA 2481
Db 2461 AAGAAGTTCAGAGCCCAAGTGA 2481
```

RESULT 8

```
AK094687
LOCUS Homo sapiens cDNA FLJ37368 fis, clone BRAMY2024530, highly similar
DEFINITION to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2
mRNA.
ACCESSION AK094687
VERSION AK094687.1 GI:21753794
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens amygdala cDNA to mRNA, clone_lib:BRAMY2
clone:BRAMY2024530.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2907)
Isogai,T. and Yamamoto,J.
Direct Submission
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..2907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAMY2024530"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="cloning vector: pME18SFL3"
BASE COUNT 696 a 788 c 839 g 584 t
ORIGIN
Query Match 94.7%; Score 2349; DB 9; Length 2907;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 5; Indels 57; Gaps 1;
Qy 1 ATGTGGGGCGCTTTGCTCGCTGCTGCGGTCCGGCGCGGACGACCATGTCGAGGAGCGC 60
Db 59 ATGTGGGGCGCTTTGCTCGCTGCTGCGGTCCGGCGCGGACGACCATGTCGAGGAGCGC 118
Qy 61 ACCATATCGCAGGACCGCGCGCGCGAGCGCGCGCAAGAGCCCGCTGCGGACCTG 120
Db 119 ACCATATCGCAGGACCGCGCGCGCGAGCGCGCGCAAGAGCCCGCTGCGGACCTG 178
Qy 121 CGCAGCGAGAGAGCGCGGACCGCTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG 180
Db 179 CGCAGCGAGAGAGCGCGGACCGCTCGGGGTGCTCCGGCGGCCCAACACCGTGTACCTG 238
Qy 181 CAGTGTGTGGCAGCGGCTAGCGGGACTCGGGCGCGCGCTCTACGTTCTTCCAGGATTC 240
Db 239 CAGTGTGTGGCAGCGGCTAGCGGGACTCG----- 268
Qy 241 AACCGGTATCTTTCAACTGTGGAGAGCGGTTTCAGAGACTCATCGAGGACCAAGTTA 300
Db 269 -----GGCGTTTCAGAGACTCATCGAGGACCAAGTTA 301
```

Qy 301 AAGGTTGCTGCGCTGGACAAATATTCTTGACAGAAATGCACTGGTCTAAATGTTGGGGGC 360
Db |||||
Qy 302 AAGGTTGCTGCGCTGGACAAATATTCTTGACAGAAATGCACTGGTCTAAATGTTGGGGGC 361
Db |||||
Qy 361 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTTCTGGA 420
Db |||||
Qy 362 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGCTACTTTCTTCTGGA 421
Db |||||
Qy 421 CTTCCAACTCGGAAATAATCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 480
Db |||||
Qy 422 CTTCCAACTCGGAAATAATCTCGAAGCAATCAAAAATATTTTCTGGTCCATTTGAAAGGA 481
Db |||||
Qy 481 ATAGAACTGGCTGTCGGGCCCACTCTGCCAGAAATACGAGGATGAACCATGACAGTT 540
Db |||||
Qy 482 ATAGAACTGGCTGTCGGGCCCACTCTGCCAGAAATACGAGGATGAACCATGACAGTT 541
Db |||||
Qy 541 TACCAGATCCCATACACAGTGAACAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 600
Db |||||
Qy 542 TACCAGATCCCATACACAGTGAACAGAGGGGAAAGCACCAACCATGGCAGAGTCCA 601
Db |||||
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db |||||
Qy 602 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCATCTTCAGACTCCGAGTTGAATGAAAT 661
Db |||||
Qy 661 GAGCCACACCTTCCACATGCTGTTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db |||||
Qy 662 GAGCCACACCTTCCACATGCTGTTAGCCAGAGAGGGGTGAGGACTCTTCCCTGGTC 721
Db |||||
Qy 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACTCTTGGTGTCTAAAGCAAG 780
Db |||||
Qy 722 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAAACTCTTGGTGTCTAAAGCAAG 781
Db |||||
Qy 781 GAGATGGGCTCCAGTGGGACAGTCCATCGCTCCCATCAATTCGCTGTCAAGAC 840
Db |||||
Qy 782 GAGATGGGCTCCAGTGGGACAGTCCATCGCTCCCATCAATTCGCTGTCAAGAC 841
Db |||||
Qy 841 GGGAAAGCATCACTCATGAAGGAGAGATTTTGGCTGAAGAGCTGTACTCTCTCA 900
Db |||||
Qy 842 GGGAAAGCATCACTCATGAAGGAGAGATTTTGGCTGAAGAGCTGTACTCTCTCA 901
Db |||||
Qy 901 GATCTGTGTGCTGCTTTTGTGTGTGTAGAAATCTCCAGATGAAAGCTTCATTCACCCATC 960
Db |||||
Qy 902 GATCTGTGTGCTGCTTTTGTGTGTGTAGAAATCTCCAGATGAAAGCTTCATTCACCCATC 961
Db |||||
Qy 961 TGTGAATGCACTTTTCAAGGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1020
Db |||||
Qy 962 TGTGAATGCACTTTTCAAGGTACCAAGAAAGGAGATGCCCGCTGGCTTGGTG 1021
Db |||||
Qy 1021 GTTCATGCGCCCGCAGCATCTGTGCTTGTGGACAGCGGTACCGAGCTGGATGGAGG 1080
Db |||||
Qy 1022 GTTCATGCGCCCGCAGCATCTGTGCTTGTGGACAGCGGTACCGAGCTGGATGGAGG 1081
Db |||||
Qy 1081 TTTTGGGCTTGACACCCAGCATCTGTGCTTGTGAATGAGAACTGTGCTCTCAGTTTCAACACTT 1140
Db |||||
Qy 1082 TTTTGGGCTTGACACCCAGCATCTGTGCTTGTGAATGAGAACTGTGCTCTCAGTTTCAACACTT 1141
Db |||||
Qy 1141 CGCAGCCAAAGATTCAAACCCAGCTCAACCTCATCCACCCGGAATCTTCCCTGCTC 1200
Db |||||
Qy 1142 CGCAGCCAAAGATTCAAACCCAGCTCAACCTCATCCACCCGGAATCTTCCCTGCTC 1201
Db |||||
Qy 1201 ACCAGTTTCCGCTGTAAGAAAGAGGGCCCACTCTGAGTGGTCCCATGGTTGAGGTTGAA 1260
Db |||||
Qy 1261 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Db |||||
Qy 1262 TGCCTCTCAAGTACAGCTCCGCTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT 1321
Db |||||
Qy 1321 TGCATCTGTAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCAATTCACAGAGAGCGTG 1380
Db |||||
Qy 1322 TGCATCTGTAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCAATTCACAGAGAGCGTG 1381
Db |||||
Qy 1381 CAGGAGTACAGGAGGTGCGCAGGACGGCCCGCCAGCAGAGAGAAAGTCAAGTAC 1440
Db |||||

Db 1382 CAGAGTACAGGAGAGTGGCAGAGCGGCCAGCCCCAGAGAGAAAGTCAAGTAC 1441
Qy |||||
Qy 1441 CCAGAAATCATCTTCTTGGAAAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1500
Db |||||
Qy 1442 CCAGAAATCATCTTCTTGGAAAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT 1501
Db |||||
Qy 1501 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGTACTTGGACTGTGGTGGGGCACA 1560
Db |||||
Qy 1502 GCCACACTTGTCAACATAAGCCCCGACACGTCTCTGTACTTGGACTGTGGTGGGGCAG 1561
Db |||||
Qy 1561 TTTGGGAGCTGTCCCTCATTCAGGAGACCAAGTGGACAGGGTCTTGGGACACCTGGCT 1620
Db |||||
Qy 1562 TTTGGGAGCTGTCCCTCATTCAGGAGACCAAGTGGACAGGGTCTTGGGACACCTGGCT 1621
Db |||||
Qy 1621 GCTGTGTTTGTGTCACCTGCGCAGATCAACACAGGGCTTGCACAGTATCTTGTCTG 1680
Db |||||
Qy 1622 GCTGTGTTTGTGTCACCTGCGCAGTCAACACAGGGCTTGCACAGTATCTTGTCTG 1681
Db |||||
Qy 1681 CAGAGAAACCGGCTTGGCATCTTGGGAAAGCGCTTCAACCTTGTCTGGTGGTGGCC 1740
Db |||||
Qy 1682 CAGAGAAACCGGCTTGGCATCTTGGGAAAGCGCTTCAACCTTGTCTGGTGGTGGCC 1741
Db |||||
Qy 1741 CCCAACAGCTCAAAGCTGSGCTCCAGCAGTACCAGAACCAAGTCCAGGAGGTCTCGCAC 1800
Db |||||
Qy 1742 CCCAACAGCTCAAAGCTGSGCTCCAGCAGTACCAGAACCAAGTCCAGGAGGTCTCGCAC 1801
Db |||||
Qy 1801 CACATCAGTATGATTCCTGCGCAATGCTTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1860
Db |||||
Qy 1802 CACATCAGTATGATTCCTGCGCAATGCTTTCAGGAGGGCTGAGATCTCCAGTCTGCA 1861
Db |||||
Qy 1861 GTGAAAGATTGATCAGTTGCTGTTCGGAACATGATGATTTGGAAGAGTTTCAGACCTGT 1920
Db |||||
Qy 1862 GTGAAAGATTGATCAGTTGCTGTTCGGAACATGATGATTTGGAAGAGTTTCAGACCTGT 1921
Db |||||
Qy 1921 CTGCTGCGGACCTCAAGCAGTGGCTTGGCTGCTGCTGGTGACACCTCTGGCTGGAAA 1980
Db |||||
Qy 1922 CTGCTGCGGACCTCAAGCAGTGGCTTGGCTGCTGCTGGTGACACCTCTGGCTGGAAA 1981
Db |||||
Qy 1981 GTGCTCTATTCGCGGACACCATGCCCTGCGAGGCTCTGCTGCGGATGGGAAAGATGCC 2040
Db |||||
Qy 1982 GTGCTCTATTCGCGGACACCATGCCCTGCGAGGCTCTGCTGCGGATGGGAAAGATGCC 2041
Db |||||
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAGATGGTTTGGAGAGGAGCAGTGGAAAG 2100
Db |||||
Qy 2042 ACCCTCTGATACATGAAGCCACCTCGAAGATGGTTTGGAGAGGAGCAGTGGAAAG 2101
Db |||||
Qy 2101 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGAGTTCAAT 2160
Db |||||
Qy 2102 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGAGTTCAAT 2161
Db |||||
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Db |||||
Qy 2162 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2221
Db |||||
Qy 2221 GAGAAAGTGGAGTTGCTTTGACCAATGAAGTCTGCTTTGAGAGACTTTCCAAACATG 2280
Db |||||
Qy 2222 GAGAAAGTGGAGTTGCTTTGACCAATGAAGTCTGCTTTGAGAGACTTTTCACAAATG 2281
Db |||||
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTGCGGACATCGAGGAGATGGAGG 2340
Db |||||
Qy 2282 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTTCTGCGGACATCGAGGAGATGGAGG 2341
Db |||||
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCCCTCTCTGTCCAGGAGCTGGCA 2400
Db |||||
Qy 2342 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGCGGCCCTCTCTGTCCAGGAGCTGGCA 2401
Db |||||
Qy 2401 GCGGCGCTTGGAGATGGGAGCTTCAGCAGAGCGGCGGCCACACAGAGGAGCCACAGGCC 2460
Db |||||
Qy 2402 GCGGCGCTTGGAGATGGGAGCTTCAGCAGAGCGGCGGCCACACAGAGGAGCCACAGGCC 2461
Db |||||
Qy 2461 AAGAAAGTCAAGAGCCCAAGTGA 2481
Db |||||


```
Db 2462 AAGAAGTTCAGAGCCCACTGA 2482
RESULT 9
AK094333
LOCUS AK094333 2734 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ37014 fis, clone BRACE2010203, highly similar
to Homo sapiens prostate cancer susceptibility protein HPC2/ELAC2
mRNA.
ACCESSION AK094333
VERSION AK094333.1 GI:21753374
KEYWORDS oligo capping; fis (full insert sequence);
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
Clone:BRACE2010203.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
AUTHORS Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Negahari, K., Masuho, Y., Negai, K. and Isogai, T.
NEO human cDNA sequencing project
Unpublished
JOURNAL 2 (bases 1 to 2734)
REFERENCE Isogai, T. and Yamamoto, J.
AUTHORS Direct Submission
TITLE Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..2734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2010203"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"
BASE COUNT 671 a 708 c 724 g 631 t
ORIGIN
Query Match 70.3%; Score 1745; DB 9; Length 2734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 737 AGCTTCACCTTAAGAGAGGAAACTCTTGCTGCTCAAAGCAAAGGAGATGGCGCTCCACG 796
Db 565 AGCTTCACCTTAAGAGAGGAAACTCTTGCTGCTCAAAGCAAAGGAGATGGCGCTCCACG 624
QY 797 TTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGACGGGAAAGATCATCTC 856
Db 625 TTGGGACAGCTGCCATCGCTCCCATCATTTGCTGCTCAAGGACGGGAAAGATCATCTC 684
QY 857 ATCAAGGAGAGAGATTTTGGCTGAGAGCTGTGTACTCTCCAGATCTGTGTGCTGCTT 916
Db 685 ATCAAGGAGAGAGATTTTGGCTGAGAGCTGTGTACTCTCCAGATCTGTGTGCTGCTT 744
QY 917 TTGTGTGTGTAGATGCTCCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAAATGCCACCT 976
Db 745 TTGTGTGTGTAGATGCTCCAGATGAAAGCTTCATTCAACCCCATCTGTGAGAAATGCCACCT 804
QY 977 TTCAGAGGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTGGTTTCATGGCCCCCAG 1036
```

Db 805 TTCAGAGGTACCAAGGAAAGGAGATGCCCGCTGGCTTGGTGGTTTCACATGGCCCCCAG 864

QY 1037 CATCTGTCTTGTGACAGCAGGTTACACAGCAGTGTGATGGAGAGGTTTGGGCTGTACACCC 1096

Db 865 CATCTGTCTTGTGACAGCAGGTTACACAGCAGTGTGATGGAGAGGTTTGGGCTGTACACCC 924

QY 1097 AGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTTCAAACTTCGACGCCACAAGATTC 1156

Db 925 AGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTTCAAACTTCGACGCCACAAGATTC 984

QY 1157 AAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTCTGCTACCACTTTCCTCTGTA 1216

Db 985 AAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTCTGCTACCACTTTCCTCTGTA 1044

QY 1217 AGAAGGAGGGCCCCACCTCAGTGTGCCCATGTGGTTTCAGGTTGAATGCCTCTCAAGTACC 1276

Db 1045 AGAAGGAGGGCCCCACCTCAGTGTGCCCATGTGGTTTCAGGTTGAATGCCTCTCAAGTACC 1104

QY 1277 AGCTCCGTCCAGGAGGAGTGGCAGAGGATGCCATTATTACTTTGCAATCTGTAGGAAT 1336

Db 1105 AGCTCCGTCCAGGAGGAGTGGCAGAGGATGCCATTATTACTTTGCAATCTGTAGGAAT 1164

QY 1337 TCATAGTTGAGCGCTGCAGCTTCCCACTTCCAGCAGCGTGCAGGAGTACAGAGGA 1396

Db 1165 TCATAGTTGAGCGCTGCAGCTTCCCACTTCCAGCAGCGTGCAGGAGTACAGAGGA 1224

QY 1397 GTGCGCAGGAGCGGCCAGCCCCAGCAGAGAAAGAGTCACTACCCAGAAATCATCTTCC 1456

Db 1225 GTGCGCAGGAGCGGCCAGCCCCAGCAGAGAAAGAGTCACTACCCAGAAATCATCTTCC 1284

QY 1457 TTGGAACAGGCTGTGCCATCCGATGAAGATTCGAAATGTTCAGTCCCACTTTGTCAACA 1516

Db 1285 TTGGAACAGGCTGTGCCATCCGATGAAGATTCGAAATGTTCAGTCCCACTTTGTCAACA 1344

QY 1517 TAAGCCCCGACACGCTCTCTGCTACTGAGTGTGTGAGGACACATTTGGGAGCTGTGCC 1576

Db 1345 TAAGCCCCGACACGCTCTCTGCTACTGAGTGTGTGAGGACACATTTGGGAGCTGTGCC 1404

QY 1577 GTCATTACGGAGACAGGTGGACAGGGTCTGGGACCCCTGCTGCTGTGTTGTGCTCC 1636

Db 1405 GTCATTACGGAGACAGGTGGACAGGGTCTGGGACCCCTGCTGCTGTGTTGTGCTCC 1464

QY 1637 ACCTGCACGACAGATCACACACGCGCTTGCCTCAAGTATCTTGTGCAGAGAGAACGCGCT 1696

Db 1465 ACCTGCACGACAGATCACACACGCGCTTGCCTCAAGTATCTTGTGCAGAGAGAACGCGCT 1524

QY 1697 TGGCATCTTTGGGAAAGCGCTTCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1756

Db 1525 TGGCATCTTTGGGAAAGCGCTTCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1584

QY 1757 CCTGGCTCCAGCAGTACCACAACAGTCCAGGAGTCTTGACCAATCATCATGATGATTC 1816

Db 1585 CCTGGCTCCAGCAGTACCACAACAGTCCAGGAGTCTTGACCAATCATCATGATGATTC 1644

QY 1817 CTGCCAAATGCTTTCAGGAAAGGCTGAGATCTCCAGTCTCTGAGTGGAAAGATGATCA 1876

Db 1645 CTGCCAAATGCTTTCAGGAAAGGCTGAGATCTCCAGTCTCTGAGTGGAAAGATGATCA 1704

QY 1877 GTTGTCTGTTGGAAACATGTGATTTGGAAGAGTTTTCAGACCTGTCTGCTGGGCACTGCA 1936

Db 1705 GTTGTCTGTTGGAAACATGTGATTTGGAAGAGTTTTCAGACCTGTCTGCTGGGCACTGCA 1764

QY 1937 AGCATGCTGTTGGCTGTGGCTGTGGTGCACCTCTGGCTGGAAAGTGTCTATTCCGGGG 1996

Db 1765 AGCATGCTGTTGGCTGTGGCTGTGGTGCACCTCTGGCTGGAAAGTGTCTATTCCGGGG 1824

QY 1997 ACACCATGCTTGCAGGCTCTGGTCCGGATGGGAAAGATGCCACCTCTCTGATACATG 2056

Db 1825 ACACCATGCTTGCAGGCTCTGGTCCGGATGGGAAAGATGCCACCTCTCTGATACATG 1884

QY 2057 AAGCCACCTGGAAGATGTTTGGAAAGGAGGAGTGGAAAGACACACAGCAACAGT 2116

```
Db 1885 AAGCCACCTCGAAGATGGTTTGGAGAGGAGCAGTGGAAAAGACACACAGCAACAGT 1944
Qy 2117 CCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAATATGCTGAACCACTTCA 2176
Db 1945 CCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAATATGCTGAACCACTTCA 2004
Qy 2177 GCCAGCGTATGCCAAGGTCCTTCTTCAGCCCCCACTTCAGCGAGAAAGTGGGAGTTG 2236
Db 2005 GCCAGCGTATGCCAAGGTCCTTCTTCAGCCCCCACTTCAGCGAGAAAGTGGGAGTTG 2064
Qy 2237 CTTTGACACATGAAGTCTGCTTTGGAGACTTTCCAACTATGCCAAGCTGATTCCTCC 2296
Db 2065 CTTTGACACATGAAGTCTGCTTTGGAGACTTTCCAACTATGCCAAGCTGATTCCTCC 2124
Qy 2297 CACTGAAGCCCTGTTGCTGCGCATCGAGGAGATGGAGGCGCAGGAGAAAGCGG 2356
Db 2125 CACTGAAGCCCTGTTGCTGCGCATCGAGGAGATGGAGGCGCAGGAGAAAGCGG 2184
Qy 2357 AGCTGCGCAGTGGCGGCGCTCTCTGTCAGGAGCTGGCAGGCGCTGGAGGATG 2416
Db 2185 AGCTGCGCAGTGGCGGCGCTCTCTGTCAGGAGCTGGCAGGCGCTGGAGGATG 2244
Qy 2417 GGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCGCAGAGGCC 2476
Db 2245 GGGAGCCTCAGCAGAAAGCGGCGCCACACAGAGAGCCACAGGCGCAGAGGCC 2304
Qy 2477 AGTGA 2481
Db 2305 AGTGA 2309

RESULT 10
AK074244
LOCUS Homo sapiens cDNA FLJ23664 fis, clone HEP03495, highly similar to
DEFINITION Putative prostate cancer susceptibility protein.
ACCESSION AK074244
VERSION AK074244.1 GI:18676795
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
clone:HEP03495.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2163)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
Location/Qualifiers
1..2163
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP03495"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
```



```

Db 2148 AATGCGGGTTCATGCTGTAACCACTTCAGTCAGCGGTACGCCAAGATCCCCCTTTTC 2207
QY 2206 AGCCCCAACTTCAGCCAGAGAAAGTGGAGTTCCTTTTGACACATGAAGTCTGCTTTTGA 2265
Db 2208 AGCCCTGACTTCAACGAGAAAGTTGCGATCGCCTTTTGACACATGAAGTCTGCTTTGA 2267
QY 2266 GACTTTCCAACTATGCCCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTTCTGCGGACATC 2325
Db 2268 GACTTCCGACAGTGCCTCAAGCTGATTTCCCTCACTGAAGCCCTGTTTGCAGGTGACAT 2327
QY 2326 GAGGAGATGGAGAGCGGAGGAGAAAGCGGAGCTCGCGGAGGTGGGGCGGCCCTTCCTG 2385
Db 2328 GAAGACATGGTGAACGCGAGGAGAGAGGGAGCTACGGCTGCTGCGAGCGCCCTCCTG 2387
QY 2386 TCCAGGAGCTGCGAGCGGCCCTGGAGGATGGGAGCCTCAGCAGAAAGCGGGCCACACA 2445
Db 2388 ACC--CAGCAGGCGAGACGCCAGAGGACAGAGAACCCCAACAGAAAGCGGGCCACACA 2444
QY 2446 GAGGAGCCACA 2456
Db 2445 GATGAACACACA 2455

RESULT 12
AC005277
LOCUS Homo sapiens cDNA FLJ36893 fis, clone UTERU2008901, highly similar
DEFINITION HPC2/ELAC2 mRNA.
ACCESSION AK094012
VERSION AK094012.1 GI:21752986
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens uterus cDNA to mRNA, clone_lib:UTERU2
clone:UTERU2008901.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takanashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2793)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UTERU2008901"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"
610 a 754 c 806 g 623 t

BASE COUNT
ORIGIN

FEATURES
source
AC005277
DEFINITION Homo sapiens chromosome 17, clone HRPK.597_M_12, complete sequence.
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 118788)

```

```

Query Match 33.2%; Score 823; DB 9; Length 2793;
Best Local Similarity 100.0%; Pred. No. 7.1e-181;
Matches 823; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1659 GGGCTTGGCAAGTATCTTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCGCT 1718
Db 1546 GGGCTTGGCAAGTATCTTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCGCT 1605
QY 1719 TCACCTTTTGTGTGTGTGGTGGCCCCCAACAGCTCAAGCTGAGCTTCCAGAGTACACAA 1778
Db 1606 TCACCTTTTGTGTGTGTGGTGGCCCCCAACAGCTCAAGCTGAGCTTCCAGAGTACACAA 1665
QY 1779 CCAGTGCCAGAGAGTCCCTGCACACATCAGTATGATTTCTGCGCAATGCCCTTCAGGAAG 1838
Db 1666 CCAGTGCCAGAGAGTCCCTGCACACATCAGTATGATTTCTGCGCAATGCCCTTCAGGAAG 1725
QY 1839 GGCTCAGATCTCCAGTCTCGAGTGGAAAGATTGATCAGTTGCTGTTGGCGATGGA 1898
Db 1726 GGCTCAGATCTCCAGTCTCGAGTGGAAAGATTGATCAGTTGCTGTTGGCGATGGA 1785
QY 1899 TTTGGAAGAGTTTCAGACCTGTCTGGTGGCGACCTGCAAGCATGCGTTTGGCTGCGCT 1958
Db 1786 TTTGGAAGAGTTTCAGACCTGTCTGGTGGCGACCTGCAAGCATGCGTTTGGCTGCGCT 1845
QY 1959 GGTGCACACCTCTGCTGGAAAGTGGTCTATTCGGGGACACCATGCCCTGCGAGGCTCT 2018
Db 1846 GGTGCACACCTCTGCTGGAAAGTGGTCTATTCGGGGACACCATGCCCTGCGAGGCTCT 1905
QY 2019 GGTCCGGATGGGAAAGATGCCACCTCTCTGATCATGAAGCCACCTTGGAAAGATGGTTT 2078
Db 1906 GGTCCGGATGGGAAAGATGCCACCTCTCTGATCATGAAGCCACCTTGGAAAGATGGTTT 1965
QY 2079 GGAAGAGGAAGCAGTGGAAAGAGACACAGCACAACGCTCCCAAGCCATCAGCGTGGGAT 2138
Db 1966 GGAAGAGGAAGCAGTGGAAAGAGACACAGCACAACGCTCCCAAGCCATCAGCGTGGGAT 2025
QY 2139 GCGGATGAACGGGAGTTCATTATGCTGAACACATTCAGCAGCGCTATGCCAAGTCCC 2198
Db 2026 GCGGATGAACGGGAGTTCATTATGCTGAACACATTCAGCAGCGCTATGCCAAGTCCC 2085
QY 2199 CCTCTTCAGCCCCCACTTCAGCGAGAAAGTGGAGTTCCTTTGACCATCATGAAGTCTG 2258
Db 2086 CCTCTTCAGCCCCCACTTCAGCGAGAAAGTGGAGTTCCTTTGACCATCATGAAGTCTG 2145
QY 2259 CTTTGGAGACTTTTCAACAATGCCCAAGCTGATTTCCCTCCTGAAAGCCCTGTTTGTCTGG 2318
Db 2146 CTTTGGAGACTTTTCAACAATGCCCAAGCTGATTTCCCTCCTGAAAGCCCTGTTTGTCTGG 2205
QY 2319 CGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGCTCGGGAGTTCGGGGCGGC 2378
Db 2206 CGACATCGAGGAGATGGAGGAGCGCAGGAGAGCGGAGTTCGGGGAGTTCGGGGCGGC 2265
QY 2379 CCTCTGTCCAGGAGCTGGCAGCGGCGCTGAGGATGGGAGCTCAGCAGAGCGGCG 2438
Db 2266 CCTCTGTCCAGGAGCTGGCAGCGGCGCTGAGGATGGGAGCTCAGCAGAGCGGCG 2325
QY 2439 CCACACAGAGGAGCCACAGGCCAAGAAGTTCAGAGCCCAAGTGA 2481
Db 2326 CCACACAGAGGAGCCACAGGCCAAGAAGTTCAGAGCCCAAGTGA 2368

RESULT 13
AC005277
LOCUS Homo sapiens chromosome 17, clone HRPK.597_M_12, complete sequence.
DEFINITION AC005277
ACCESSION AC005277
VERSION AC005277.1 GI:3337311
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 118788)

```



```

repeat_region      /rpt_family="LiM4"
complement(21071..21100)
/rpt_family="(TAAA)n"
repeat_region      complement(21101..21338)
/rpt_family="AluSx"
repeat_region      complement(21348..21598)
/rpt_family="LiM4"
repeat_region      21799..22104
/rpt_family="AluJo"
repeat_region      22151..22562
/rpt_family="(MSTA"
repeat_region      23978..24287
/rpt_family="AluSq"
repeat_region      24852..25199
/rpt_family="L2"
repeat_region      complement(25697..25985)
/rpt_family="MLT1b"
unsure             25816..25895
/note="Single-stranded coverage."
repeat_region      complement(25994..26099)
/rpt_family="purine-rich"
repeat_region      complement(26122..26291)
/rpt_family="MLT1b"
repeat_region      26312..26337
/rpt_family="(CAAAA)n"
repeat_region      26456..26480
/rpt_family="(GGGAA)n"
repeat_region      complement(26576..26781)
/rpt_family="MER6B"
repeat_region      complement(26912..26990)
/rpt_family="MER21B"
repeat_region      26991..27290
/rpt_family="AluSx"
repeat_region      complement(27291..27975)
/rpt_family="MER21B"
repeat_region      28034..28197
/rpt_family="LTR28"
repeat_region      28266..28795
/rpt_family="LTR28"
repeat_region      28843..28916
/rpt_family="LTR28"
repeat_region      28920..29005
/rpt_family="MIR"
repeat_region      29140..29233

Query Match      10.0%; Score 247.4; DB 9; Length 118788;
Best Local Similarity 97.7%; Pred. No. 1.1e-46;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGTGGCGCTTTGCTGCTGTGCGGTCCGGCGGCGGACGACCATGTGCGAGGACGC 60
Db 78562 ATGTGGCGCTTTGCTGCTGTGCGGTCCGGCGGCGGACGACCATGTGCGAGGACGC 78621

QY 61 ACCATATCCAGGACCCCGCCCGCGGAGCGCGCGAAGGACCGCTGCGGACCTG 120
Db 78622 ACCATATCCAGGACCCCGCCCGCGGAGCGCGCGAAGGACCGCTGCGGACCTG 78681

QY 121 CGCACGCGAGAGAAGCGCGGACCGTCCGGGTGCTCCGGCGGCGGACACCGGTGTACCTG 180
Db 78682 CGCACGCGAGAGAGCGCGGACCGTCCGGGTGCTCCGGCGGCGGACACCGGTGTACCTG 78741

QY 181 CAGGTGGTGGCAGCGGAGTACCGGGAGCTCGGGCGCGCGCTCTACGTCCTTCCGAGTTC 240
Db 78742 CAGGTGGTGGCAGCGGAGTACCGGGAGCTCGGGCGCGCGCTCTACGTCCTTCCGAGTTC 78801

QY 241 AACCGTATCTCTTCAA 257
Db 78802 AACCGTATCTCTTCAA 78818

RESULT 14
HSA329331
LOCUS              HSA329331          679 bp      DNA      linear      PRI 18-JUL-2002

DEFINITION      Homo sapiens genomic sequence surrounding Not1 site, clone
                NL6-B020C.
ACCESSION      AJ329331
VERSION        AJ329331.1
KEYWORDS       GI:15873749
SOURCE         Homo sapiens.
ORGANISM       Homo sapiens.
REFERENCE      1 (bases 1 to 679)
AUTHORS        Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
                Podowski,R.M., Matushkin,Y.G., Gvanchandani,A., Muravenko,O.V.,
                Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I.,
                Kiselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE          Not1 flanking sequences: a tool for gene discovery and verification
                of the human genome
JOURNAL        Nucleic Acids Res. 30 (14), 3163-3170 (2002)
PUBMED        12136098
REFERENCE      2 (bases 1 to 679)
AUTHORS        Zabarovsky,E.R.
TITLE          Direct Submission
JOURNAL        Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
                Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
                Sweden
FEATURES        Location/Qualifiers
                source          1..679
                /Organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="NL6-B020C"
BASE COUNT      123 a 172 c 179 g 205 t
ORIGIN          1
                6.3%; Score 156.8; DB 9; Length 679;
                Best Local Similarity 95.8%; Pred. No. 1.1e-25;
                Matches 161; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 90 GCGGCGCGCAAGGACCGCTGCGGCACCTGCGGACGAGAGAGCGGACCGCTCGGG 149
Db 1 GCGGCGCGCAAGGACCGCTGCGGCACCTGCGGACGAGAGAGCGGACCGCTCGGG 60

QY 150 GTGCTCCGCGCGCCCAACACCGTGTACTGTCAGTGTGGCAGCGGTAGCCGGAGCTC 209
Db 61 GTGCTCCGCGCGCCCAACACCGTGTACTGTCAGTGTGGCAGCGGTAGCCGGAGCTC 120

QY 210 GCGGCGCGCGCTCTACGTCCTTCTCCGAGTTCAACCGGTATCTTTCAA 257
Db 121 GCGGCGCGCGCTCTACGTCCTTCTCCGAGTTCAACCGGTCAAGTCAACGA 168

RESULT 15
AF304371S2
LOCUS              AF304371S2          740 bp      DNA      linear      PRI 19-OCT-2000
DEFINITION      Homo sapiens putative prostate cancer susceptibility protein
                HPC2/BLAC2 gene, exon 17 and partial cds.
ACCESSION      AF304369
VERSION        AF304369.1
KEYWORDS       GI:10880929
SEGMENT        2 of 2
SOURCE         Homo sapiens.
ORGANISM       Homo sapiens.
REFERENCE      1 (bases 1 to 740)
AUTHORS        Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
                Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
                Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
                Gupta,J.S., Hu,R., Iliev,D., Janeczek,T., Kort,E.N., Laity,K.E.,
                Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
                Peterson,K.T., Reid,J.B., Richards,S., Schroeder,M., Smith,R.,
                Snyder,S.C., Swedlund,B., Swensen,J., Thomas,A., Tranchant,M.,
                Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.,
                and Cannon-Albright,L.A.
TITLE          A strong candidate prostate cancer susceptibility gene at

```


chromosome 17p
 Unpublished
 2 (bases 1 to 740)
 JOURNAL
 REFERENCE
 AUTHORS
 Tavtigian,S.V., Simard,J., Teng,D.H.F., Baumgard,M., Beck,A.,
 Camp,N.J., Carillo,A.R., Chen,Y., Dayananth,P., Desrochers,M.,
 Dumont,M., Farnham,J.M., Frank,D., Frye,C., Ghaffari,S.,
 Gupte,J.S., Hu,R., Iliev,D., Janeski,T., Kort,E.N., Laity,K.E.,
 Leavitt,A., Leblanc,G., McArthur-Morrison,J., Pederson,A., Penn,B.,
 Peterson,K.T., Reid,J.E., Richards,S., Schroeder,M., Smith,R.,
 Snyder,S.C., Swelund,B., Swensen,J., Thomas,A., Tranchant,M.,
 Woodland,A.-M., Labrie,F., Skolnick,M.H., Neuhausen,S., Rommens,J.
 and Cannon-Albright,L.A.
 TITLE
 JOURNAL
 FEATURES
 source
 Direct Submission
 Submitted (12-SEP-2000) Myriad Genetics, Inc., 320 Wakara Way, Salt
 Lake City, UT 84108, USA
 Location/Qualifiers
 1..740
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17p"
 mRNA
 <301..>439
 /product="putative prostate cancer susceptibility protein
 HPC2/ELAC2"
 CDS
 <301..>439
 /codon_start=2
 /product="putative prostate cancer susceptibility protein
 HPC2/ELAC2"
 /protein_id="AAG24440.1"
 /db_xref="GI:10880931"
 /translation="PDTSLLLDCGEGTFGLCRHYGQVDRVLGTLAAVFVSHLHADH
 HT"
 exon
 301..439
 /number=17
 BASE COUNT 154 a 188 c 173 g 225 t
 ORIGIN
 Query Match 5.9%; Score 146.2; DB 9; Length 740;
 Best Local Similarity 89.7%; Pred. No. 3.3e-23;
 Matches 157; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 1486 ATTCGAATGTGAGTCCACACTTGTCAACATAAGCCCGACACGCTCTCTGCTACTGGAC 1545
 DB 266 AGTGGCAGTGATCTCTTCTCTCTCTGCGCCCGACACGCTCTCTGCTACTGGAC 325
 QY 1546 TGTGTGAGGGCACATTTGGGCGAGCTGTGCCCTCATTTACGGAGACAGGTGCACAGGGTC 1605
 DB 326 TGTGTGTAGGGCACATTTGGGCGAGCTGTGCCCTCATTTACGGAGACAGGTGCACAGGGTC 385
 QY 1606 CTGGGCACCTGGCTGTGTGTTGTGTCCTCCACCTGCAGCGAGATCACACACGG 1660
 DB 386 CTGGGCACCTGGCTGTGTGTTGTGTCCTCCACCTGCAGCGAGATCACACACGG 440
 Search completed: May 17, 2003, 21:25:41
 Job time : 5600.12 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2481	100.0	2481	4	US-09-564-805-1	Sequence 1, Appli
2	2481	100.0	2958	4	US-09-564-805-3	Sequence 3, Appli
3	2455.4	99.0	2908	4	US-09-564-805-223	Sequence 223, App
4	2442.6	98.5	2892	4	US-09-564-805-225	Sequence 225, App
5	1645.6	66.3	2470	4	US-09-564-805-221	Sequence 221, App
6	247.4	10.0	350	4	US-09-564-805-210	Sequence 210, App
7	247.4	10.0	26664	4	US-09-564-805-28	Sequence 28, Appli
8	245	9.9	295	4	US-09-564-805-4	Sequence 4, Appli
9	237	9.6	238	4	US-09-328-111-315	Sequence 315, App
10	228	9.2	655	4	US-09-564-805-27	Sequence 27, Appl
11	145	5.8	145	4	US-09-564-805-26	Sequence 26, Appl
12	139	5.6	139	4	US-09-564-805-16	Sequence 16, Appl
13	139	5.6	139	4	US-09-564-805-20	Sequence 20, Appl
14	121	4.9	121	4	US-09-564-805-24	Sequence 24, Appl
15	120	4.8	120	4	US-09-564-805-10	Sequence 10, Appl
16	119	4.8	119	4	US-09-564-805-18	Sequence 18, Appl
17	113.6	4.6	326	4	US-09-564-805-212	Sequence 212, App
18	113	4.6	113	4	US-09-564-805-14	Sequence 14, Appl
19	110	4.4	110	4	US-09-564-805-22	Sequence 22, Appl
20	100	4.0	100	4	US-09-564-805-23	Sequence 23, Appl
21	97	3.9	97	4	US-09-564-805-19	Sequence 19, Appl
22	96	3.9	96	4	US-09-564-805-15	Sequence 15, Appl
23	86	3.5	86	4	US-09-564-805-17	Sequence 17, Appl
24	79	3.2	79	4	US-09-564-805-25	Sequence 25, Appl
25	73	2.9	73	4	US-09-564-805-13	Sequence 13, Appl
26	71	2.9	71	4	US-09-564-805-6	Sequence 6, Appli
27	69	2.8	69	4	US-09-564-805-9	Sequence 9, Appli

Qy 241 AACCGTATCTCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGTCAGGAGCACAAGTTA 300
Db |||||
Qy 241 AACCGTATCTCTTCAACTGTGGAGAGCGCTTCAGAGACTCATGTCAGGAGCACAAGTTA 300
Db |||||
Qy 301 AAGGTTGCTCGCCTGGCAACAATATTCCTGACAGAAATGCACTGGTCTTAATGTTGGGGC 360
Db |||||
Qy 301 AAGGTTGCTCGCCTGGCAACAATATTCCTGACAGAAATGCACTGGTCTTAATGTTGGGGC 360
Db |||||
Qy 361 TTAAGTGAATGATCTTCTTAAAGAAACCGGGCTTCCAAAGTGTAATCTTCTGGA 420
Db |||||
Qy 361 TTAAGTGAATGATCTTCTTAAAGAAACCGGGCTTCCAAAGTGTAATCTTCTGGA 420
Db |||||
Qy 421 CCTCCAACTCGGAAATACCTCGAAGCAATCAAAATATTTCTGTCATTTGAAGGA 480
Db |||||
Qy 421 CCTCCAACTCGGAAATACCTCGAAGCAATCAAAATATTTCTGTCATTTGAAGGA 480
Db |||||
Qy 481 ATAGAACTGGCTGTGCGGCCCTCACTCTCCCGCAGAAATACGAGGATGAAACCATGACAGTT 540
Db |||||
Qy 481 ATAGAACTGGCTGTGCGGCCCTCACTCTCCCGCAGAAATACGAGGATGAAACCATGACAGTT 540
Db |||||
Qy 541 TACAGATCCCATACACAGTGAACAGAGGAGGGAAGCAACCAATGTCAGAGTCCA 600
Db |||||
Qy 541 TACAGATCCCATACACAGTGAACAGAGGAGGGAAGCAACCAATGTCAGAGTCCA 600
Db |||||
Qy 601 GAAAGGCTCTCAGCAGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCAATGAAAT 660
Db |||||
Qy 601 GAAAGGCTCTCAGCAGCTCAGTCCAGAGCATCTTCAGACTCCGAGTCAATGAAAT 660
Db |||||
Qy 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db |||||
Qy 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
Db |||||
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTCTCAAAAGCAAG 780
Db |||||
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTCTCAAAAGCAAG 780
Db |||||
Qy 781 GAGATGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGAC 840
Db |||||
Qy 781 GAGATGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGAC 840
Db |||||
Qy 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCA 900
Db |||||
Qy 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCA 900
Db |||||
Qy 901 GATCCTGTGTCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCACCCATC 960
Db |||||
Qy 901 GATCCTGTGTCTTTTGTGGTGTAGAAATGTCCAGATGAAAGCTTCATTCACCCATC 960
Db |||||
Qy 961 TGTGAGATGCCACTTTTCAGAGGTACCAAGAAAGGCAGATGCCCGCTGGCTTGGTG 1020
Db |||||
Qy 961 TGTGAGATGCCACTTTTCAGAGGTACCAAGAAAGGCAGATGCCCGCTGGCTTGGTG 1020
Db |||||
Qy 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGGACAGAGGTACCAAGAGTGGATGGAGAG 1080
Db |||||
Qy 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGGACAGAGGTACCAAGAGTGGATGGAGAG 1080
Db |||||
Qy 1081 TTTGGGCTGACACCCAGCATCTGTCTCTGAATGAGAACTGTGCTCAGTTTCAACACTT 1140
Db |||||
Qy 1141 GCGAGCCACAGATTCAAACCCAGCTCAACCTCATCTCCACCGGACATCTTCCCTGGTC 1200
Db |||||
Qy 1141 GCGAGCCACAGATTCAAACCCAGCTCAACCTCATCTCCACCGGACATCTTCCCTGGTC 1200
Db |||||
Qy 1201 ACCAGTTTCCGCTGTAAAGAGAGGGCCCACTTCAAGTGTGCCCATGGTTGAGGGTGA 1260
Db |||||
Qy 1201 ACCAGTTTCCGCTGTAAAGAGAGGGCCCACTTCAAGTGTGCCCATGGTTGAGGGTGA 1260
Db |||||
Qy 1261 TGCTCTCAAGTACCAAGTCCGTCCTCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db |||||
Qy 1261 TGCTCTCAAGTACCAAGTCCGTCCTCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACT 1320
Db |||||
Qy 1321 TGCAATCCTGTAGGAATTCATAGTTGAGGGCTGTCAGCTTCCCACTTCAGCAGAGCGTG 1380
Db |||||

Db 1321 TGCAATCCTGTAGGAATTCATAGTTGAGCGCTGCAGCTTCCCACTTCAGCAGAGCGTG 1380
Qy |||||
Qy 1381 CAGAGTACAGGAGGAGTGCAGGAGCGGCCAGCCCCAGCAGAGAGAGAAAGTCAGTAC 1440
Db |||||
Qy 1381 CAGAGTACAGGAGGAGTGCAGGAGCGGCCAGCCCCAGCAGAGAGAGAAAGTCAGTAC 1440
Db |||||
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAGATTCCGAAATGTCA 1500
Db |||||
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAGATTCCGAAATGTCA 1500
Db |||||
Qy 1501 GCCACACTTGTCAACATAAGCCCCGACACCTCTCTGTACTTGGACTGTGGTACGGGACA 1560
Db |||||
Qy 1501 GCCACACTTGTCAACATAAGCCCCGACACCTCTCTGTACTTGGACTGTGGTACGGGACA 1560
Db |||||
Qy 1561 TTTGGGAGCTGTGCCGTCAATTACGGAGACAAGGTGACAGGGTCTTGGGACCCCTGGCT 1620
Db |||||
Qy 1561 TTTGGGAGCTGTGCCGTCAATTACGGAGACAAGGTGACAGGGTCTTGGGACCCCTGGCT 1620
Db |||||
Qy 1621 GCTGTGTTGTGTCACCTGTCAGCGAGATCAACACAGGGCTTGGCCAAAGTATCTTGTG 1680
Db |||||
Qy 1621 GCTGTGTTGTGTCACCTGTCAGCGAGATCAACACAGGGCTTGGCCAAAGTATCTTGTG 1680
Db |||||
Qy 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGTGTC 1740
Db |||||
Qy 1681 CAGAGAAACGCGCTTGGCATCTTTGGGAAAGCGCTTCAACCTTTGCTGGTGTGTC 1740
Db |||||
Qy 1741 CCCAACAGCTCAAAGCTTGGCTCCAGCAGTACCAACACAGTGCAGGAGGTCTCTGCAC 1800
Db |||||
Qy 1741 CCCAACAGCTCAAAGCTTGGCTCCAGCAGTACCAACACAGTGCAGGAGGTCTCTGCAC 1800
Db |||||
Qy 1801 CACATCAGTATGATCTTGCCTGCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
Db |||||
Qy 1801 CACATCAGTATGATCTTGCCTGCAATGCTTTCAGGAAGGGCTGAGATCTCCAGTCTGCA 1860
Db |||||
Qy 1861 GTGGAAAGATTGATCAGTTCGCTGTTCCGAAACATGTATTTGGAAGAGTTTTCAGACCTGT 1920
Db |||||
Qy 1861 GTGGAAAGATTGATCAGTTCGCTGTTCCGAAACATGTATTTGGAAGAGTTTTCAGACCTGT 1920
Db |||||
Qy 1921 CTGTCGCGCACTGCAAGCATGCTTTGGCTGTCGCTGTCGACACCTCTGGCTGGAAA 1980
Db |||||
Qy 1921 CTGTCGCGCACTGCAAGCATGCTTTGGCTGTCGCTGTCGACACCTCTGGCTGGAAA 1980
Db |||||
Qy 1981 GTGTCCTATTCCGGGGACACATGCTCCCTCGAGGCTCTGTCGCGATGGGAAAGATGCC 2040
Db |||||
Qy 1981 GTGTCCTATTCCGGGGACACATGCTCCCTCGAGGCTCTGTCGCGATGGGAAAGATGCC 2040
Db |||||
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAAGATGCTTTGGAAGAGGAGCAGTGGAAAAG 2100
Db |||||
Qy 2041 ACCCTCTGATACATGAAGCCACCTCGAAAGATGCTTTGGAAGAGGAGCAGTGGAAAAG 2100
Db |||||
Qy 2101 ACACAGCACAACGTCCTCCAAAGCCATCAGCGTGGGGATGCGGATGAAACGCGAGTTCA 2160
Db |||||
Qy 2101 ACACAGCACAACGTCCTCCAAAGCCATCAGCGTGGGGATGCGGATGAAACGCGAGTTCA 2160
Db |||||
Qy 2161 ATGCTGAACCATCTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAGC 2220
Db |||||
Qy 2161 ATGCTGAACCATCTTCAGCAGCGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAGC 2220
Db |||||
Qy 2221 GAGAAAGTGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAAACATG 2280
Db |||||
Qy 2221 GAGAAAGTGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAAACATG 2280
Db |||||
Qy 2281 CCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTTGTGTCGCGACATCGAGGAGATCGAGAG 2340
Db |||||
Qy 2281 CCCAAGCTGATTTCCCCCACTGAAAGCCCTGTTTGTGTCGCGACATCGAGGAGATCGAGAG 2340
Db |||||
Qy 2341 GCGAGGAGAGAGCGGGAGCTGCGGACAGTGGCGGCGCCCTCTCTGTCAGGGAGCTGGCA 2400
Db |||||
Qy 2341 GCGAGGAGAGAGCGGGAGCTGCGGACAGTGGCGGCGCCCTCTCTGTCAGGGAGCTGGCA 2400
Db |||||
Qy 2401 GCGGCTGTGAGGATGGGAGCTTCAGCAGAAAGCGGCGCCACACAGAGGAGGCCACAGGCC 2460
Db |||||

```

Db   2401 GGCGGCTCGAGATGGGAGCCTACAGAGAAGCGGCCCAACAGAGAGCCACAGGCC 2460
Qy   2461 AAGAAAGTCAGAGCCCAGTGA 2481
      |||||
Db   2461 AAGAAAGTCAGAGCCCAGTGA 2481

RESULT 2
US-09-564-805-3
; Sequence 3, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.P.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-564-805-3

Query Match          100.0%; Score 2481; DB 4; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy   1 ATGTGGCGCTTTTGCTCGCTGCTGCCTCGCGCGCGGACGACCATGTCGAGGAGCG 60
Db   51 ATGTGGCGCTTTTGCTCGCTGCTGCCTCGCGCGCGGACGACCATGTCGAGGAGCG 110
Qy   61 ACCATATCGAGGACACCGCCCGCGAGCGCGCGGACGACCGCTCGGGACCCTG 120
Db   111 ACCATATCGAGGACACCGCCCGCGAGCGCGCGGACGACCGCTCGGGACCCTG 170
Qy   121 CGCACGGGAGAGAAGCGCGGACCGTCTCGGGGTGCTCGGGCGGCCAACAACACGTCGTACCTG 180
Db   171 CGCACGGGAGAGAAGCGCGGACCGTCTCGGGGTGCTCGGGCGGCCAACAACACGTCGTACCTG 230
Qy   181 CAGTGTGTGACGCGGTAGCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
Db   231 CAGTGTGTGACGCGGTAGCGGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 290
Qy   241 AACCGGTATCTTCAACTGTGTGAGAAGCGGTTTAGAGACTCATCGAGGAGACAAGTTA 300
Db   291 AACCGGTATCTTCAACTGTGTGAGAAGCGGTTTAGAGACTCATCGAGGAGACAAGTTA 350
Qy   301 AAGTTGCTCGCTGGACAAATATTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 360
Db   351 AAGTTGCTCGCTGGACAAATATTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 410
Qy   361 TTAAGTGAATGATTCTTACTTTAAAGGAAAACCGGGCTTCCAAGTGTGTACTTCTCGGA 420
Db   411 TTAAGTGAATGATTCTTACTTTAAAGGAAAACCGGGCTTCCAAGTGTGTACTTCTCGGA 470
Qy   421 CCTCCAACAATCGAAAAATACTCGAAGCAATCAAATAATTTTCTGGTCCATTGAAGGA 480
Db   471 CCTCCAACAATCGAAAAATACTCGAAGCAATCAAATAATTTTCTGGTCCATTGAAGGA 530

```

Qy	1561	TTTGGG	CAGCTGTG	CGCTCATTTAC	GAGACAC	CAGGTG	GACAGGGT	CTCTGG	GACACCTCGCT	1620
Db	1611	TTTGGG	CAGCTGTG	CGCTCATTTAC	GAGACAC	CAGGTG	GACAGGGT	CTCTGG	GACACCTCGCT	1670
Qy	1621	GC	TGCTTTT	TGTGCC	CACCTG	CACG	CAGATAC	CA	CGGCTTGC	1680
Db	1671	GC	TGCTTTT	TGTGCC	CACCTG	CACG	CAGATAC	CA	CGGCTTGC	1730
Qy	1681	CAG	GAGAA	CGCGCTT	TGGCATCTTT	TGGG	AAAGCCGCTT	CAC	CCCTTGTG	1740
Db	1731	CAG	GAGAA	CGCGCTT	TGGCATCTTT	TGGG	AAAGCCGCTT	CAC	CCCTTGTG	1790
Qy	1741	CC	CAAC	CAGCT	CAAG	CGTGG	CTCAG	CAGTAC	CA	1800
Db	1791	CC	CAAC	CAGCT	CAAG	CGTGG	CTCAG	CAGTAC	CA	1850
Qy	1801	CAC	ATCAGT	TGATTC	TCC	CAAA	TCCCTT	CAG	GAAGGG	1860
Db	1851	CAC	ATCAGT	TGATTC	TCC	CAAA	TCCCTT	CAG	GAAGGG	1910
Qy	1861	GT	GAAAG	ATTG	ATCAGTTC	CGCTT	TGCG	AA	CA	1920
Db	1911	GT	GAAAG	ATTG	ATCAGTTC	CGCTT	TGCG	AA	CA	1970
Qy	1921	CT	GTG	CG	CAC	TG	CAAG	CA	TG	2000
Db	1971	CT	GTG	CG	CAC	TG	CAAG	CA	TG	2030
Qy	1981	GT	GGT	CTATTT	CCGG	GACAC	CA	TG	CC	2040
Db	2031	GT	GGT	CTATTT	CCGG	GACAC	CA	TG	CC	2090
Qy	2041	ACC	CTCT	GTAT	ATAC	TGA	AGCC	CCCTT	GGA	2100
Db	2091	ACC	CTCT	GTAT	ATAC	TGA	AGCC	CCCTT	GGA	2150
Qy	2101	AC	AC	CAG	CA	AA	CGTCC	CA	AG	2160
Db	2151	AC	AC	CAG	CA	AA	CGTCC	CA	AG	2210
Qy	2161	AT	GT	GTAA	CCAC	TTAC	CG	CG	CTAT	2220
Db	2211	AT	GT	GTAA	CCAC	TTAC	CG	CG	CTAT	2270
Qy	2221	G	A	AA	AGTGG	AGTGG	CTG	CTT	TG	2280
Db	2271	G	A	AA	AGTGG	AGTGG	CTG	CTT	TG	2330
Qy	2281	CC	CA	AG	CTG	AT	CCCC	CA	TG	2340
Db	2331	CC	CA	AG	CTG	AT	CCCC	CA	TG	2390
Qy	2341	CG	CA	AG	GAG	AA	GG	AG	CTG	2400
Db	2391	CG	CA	AG	GAG	AA	GG	AG	CTG	2450
Qy	2401	GG	CG	GC	CT	TG	GAG	AT	GG	2460
Db	2451	GG	CG	GC	CT	TG	GAG	AT	GG	2510
Qy	2461	A	A	AG	AG	GT	C	A	G	2481
Db	2511	A	A	AG	AG	GT	C	A	G	2531

RESULT 3

RESUL 3
US-09-564-805-223
; Sequence 223, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques

Qy 721 GTAGCTTTTCATCTGTGAAGCTTCACTTAAAGAGAGAAACTTTTGGTGCTCAAGCAAG 780
Db 721 GTAGCTTTTCATCTGTGAAGCTTCACTTAAAGAGAGAAACTTTTGGTGCTCAAGCAAG 780
Qy 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGTCCCATCATCTGCTGTCAAGGAC 840
Db 781 GAGATGGGCTCCAGTTGGGACAGCTGCCATCGTCCCATCATCTGCTGTCAAGGAC 840
Qy 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 900
Db 841 GGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTTCCA 900
Qy 901 GATCCTGTGTCTGCTTTTGTGGTGTGAATGTCCAGATGAAGCTTCATTCAACCCATC 960
Db 901 GATCCTGTGTCTGCTTTTGTGGTGTGAATGTCCAGATGAAGCTTCATTCAACCCATC 960
Qy 961 TGTGAGATGTCACCTTTTCAGAGTACCAAGAAAGGAGAGATGCCCGTGGCTTGGT 1020
Db 961 TGTGAGATGTCACCTTTTCAGAGTACCAAGAAAGGAGAGATGCCCGTGGCTTGGT 1020
Qy 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGGACAGAGTACCAGAGTGGATGGAGAGG 1080
Db 1021 GTTCACATGGCCCAAGCATCTGTGCTGTGGACAGAGTACCAGAGTGGATGGAGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGCATCTGTCTCTGAATGAGAACTGTGCTCAGTTTCAAACTT 1140
Db 1081 TTTGGGCTGACACCCAGCATCTGTCTCTGAATGAGAACTGTGCTCAGTTTCAAACTT 1140
Qy 1141 CGAGCCCAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
Db 1141 CGAGCCCAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTTCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCAACCTCAGTGTGCCATGTTTCAGGGTGA 1260
Db 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCAACCTCAGTGTGCCATGTTTCAGGGTGA 1260
Qy 1261 TGCCTCTCTCAAGTACCACTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCTCAAGTACCACTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGGCTGAGCTTCCCACTTCAGCAGAGCTG 1380
Db 1321 TGCAATCTGAGGAATTCATAGTTGAGGGCTGAGCTTCCCACTTCAGCAGAGCTG 1380
Qy 1381 CAGGAGTACAGGAGAGTGGCAGGACGCGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Db 1381 CAGGAGTACAGGAGAGTGGCAGGACGCGCCAGCCAGCAGAGAGAAAGTCAAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGGAAGATTGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGGTCTGCCATCCCGATGGAAGATTGAAATGTCAGT 1500
Qy 1501 GCCACATCTGTCAATGAAGCCCGACAGCTCTGTCTGCTACTGGAATGTGGTGGAGGACA 1560
Db 1501 GCCACATCTGTCAATGAAGCCCGACAGCTCTGTCTGCTACTGGAATGTGGTGGAGGACG 1560
Qy 1561 TTTGGGAGCTGTGCGCTCATTTACGAGACAGGAGTGGACAGGCTCTGGCACCCCTGGCT 1620
Db 1561 TTTGGGAGCTGTGCGCTCATTTACGAGACAGGAGTGGACAGGCTCTGGCACCCCTGGCT 1620
Qy 1621 GCTGTGTTTGTGTCCTCACTGACGAGATCACACAGGGCTTCCCAAGTATCTTGTGTG 1680
Db 1621 GCTGTGTTTGTGTCCTCACTGACGAGATCACACAGGGCTTCCCAAGTATCTTGTGTG 1680
Qy 1681 CAGAGAGAACCGCTTGTGCTATTTTGGAAAGCCGCTTACCCCTTGTGGTGTGTGCTG 1740
Db 1681 CAGAGAGAACCGCTTGTGCTATTTTGGAAAGCCCTTACCCCTTGTGGTGTGTGCTG 1740
Qy 1741 CCCAACCCAGCTCAAGGCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGCTCTGCAC 1800
Db 1741 CCCAACCCAGCTCAAGGCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGCTCTGCAC 1800

Qy 1801 CACATCAGTATGATTCCTGCAAAATGCTTCAGGAAGGGGCTGAGATCTCCAGTCTCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCAAAATGCTTCAGGAAGGGGCTGAGATCTCCAGTCTCTGCA 1860
Qy 1861 GTGGAAGATTGATCAGTTCGCTGTCGGAACATGTGATTTGGAAAGAGTTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTGATCAGTTCGCTGTCGGAACATGTGATTTGGAAAGAGTTTTCAGACCTGT 1920
Qy 1921 CTGGTGGGCACTGCAAGCATGCTTTGGCTGTGCTGTGTCACACCTCTGGCTGGAAA 1980
Db 1921 CTGGTGGGCACTGCAAGCATGCTTTGGCTGTGCTGTGTCACACCTCTGGCTGGAAA 1980
Qy 1981 GTGCTCTATTTCGGGGACACCATGCTCCGAGGCTCTGCTCCGATGGGAAAGATGCC 2040
Db 1981 GTGCTCTATTTCGGGGACACCATGCTCCGAGGCTCTGCTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCACCCTCGAAGATGTTTGGAAAGAGAGTGGAAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCACCCTCGAAGAGCGTTTGGNAGAGGAGTGGAAAAG 2100
Qy 2101 ACACAGCACAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2160
Db 2101 ACACAGCACAACGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCATTTACGAGCAGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAAC 2220
Db 2161 ATGCTGAACCATTTACGAGCAGCTATGCCAAGTCCCTCTTCAGCCCCCACTTCAAC 2220
Qy 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAAAGCCTGTTTCTGGCGACATCGAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAAAGCCTGTTTCTGGCGACATCGAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGGGCGCTCTCTGTCAGGAGGCTGCA 2400
Db 2341 CGCAGGAGAAAGCGGAGCTCGGCGAGTGGCGGGGCGCTCTCTGTCAGGAGGCTGCA 2400
Qy 2401 GCGGCTGGAGGATGGGAGCCTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2460
Db 2401 GCGGCTGGAGGATGGGAGCCTCAGCAGAGCGGCGCCACACAGAGAGCCACAGGCC 2460
Qy 2461 AAGAAGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAGTTCAGAGCCCACTGA 2481

RESULT 4

US-09-564-805-225
; Sequence 225, Application US/09564805
; Patent No. 633403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR FILING DATE: 1998-11-06
; PRIOR FILING DATE: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA

D	b	100	CTCGCTACGCGGGAGAAAGCGCGCGCCC-----GGGTCCCGGGGGCCCGAAACACCGCTGTATC	153
Q	y	178	CTCAGCTGGTGGCAGCGGGTAGCCGGACTCGGCGCGCGGCTCTACGTCTTCTCCGAG	237
D	b	154	CTCAGCTGGTGGCGGGCGGGACCGGGGGCTGCTCTATGCTTCTCGGAA	213
Q	y	238	TTCAACGGGTATCTCTTCAACTGTGGAGAGGGGTTGAGAGATCATGCAAGGAGCAACAG	297
D	b	214	TACAACAGGTACCTTTTAACTTCGCGGAGAGCGCTCCAACGACTTATGCAAGAACACAAG	273
Q	y	298	TTAAAGGTTGCTCGCTCGACACAACATATCTCTGACACGAATGCACCTGGTCTAAATGTTGG	357
D	b	274	ACTGAAGTCGCTCGCTTGACAAACATCTTCTGACTCGGATGCAATGGTCAAAATGTTGG	333
Q	y	358	GGCTTAAAGTCGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGACTTTCT	417
D	b	334	GGGTTGTGTGGAATGATTTTAACTTTAAAGGAACCGGGCTTCCCAAATGTGTTCTGTCT	393
Q	y	418	GGACTTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTTCTGGTCCATTTGAA	477
D	b	394	GGACCACCAAGCTGGAGAAATATCTAGAAGCAATCAAAATATTTTTCTGGTCCATTTGAA	453
Q	y	478	GGATAGAACTGGCTGTGCGGCCCACTCTGCCACGAATACGAGGATGAACCATGACA	537
D	b	454	GGATAGAACTGGCGTGGCGCTCCTCTCTGCACAGAAATCAAGGATGAGACCATGACT	513
Q	y	538	GTTTTACAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCAACCATGGCAGGT	597
D	b	514	GTTTACAGGTCCCTATCCACAGTGAACGGAGGTGTGGAAAGCAACAGCCATCCCAGAGC	573
Q	y	598	CCAGAAGGCCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCAATGAA	657
D	b	574	CCAGAAACATCTCCCAACAGGCTCAGTCCCAAACAGTCACTCGGACTCTGGATCAGCTGA	633
Q	y	658	AATGAGCCACCTTCCACATGCTGTAGCCAGAGAGAGGGGTACGGGACTCTTCCCTG	717
D	b	634	AATGGGC-----AGTCCCAACAGGAAGCATGGGGCAGGGAC-CCTCCTTA	678
Q	y	718	GTCTGTAGCTTTTCATCTGTAAAGCTTCACCTTAAAGAGAGAAACTTTCTGGTGTCAAGCA	777
D	b	679	GTGGTAGCTTTTGTCTGCAAGCTTCACCTTGAGGAAGGAAACTTCTTGGTGTCTTAAAGCA	738
Q	y	778	AAGGAGATGGGCTCCAGTTGGGACAGCTCCCATCGCTCCCATCATTTGCTGTGTCAAG	837
D	b	739	AAGGAGCTGGGCTTCTCTGTTGGGACGCGCCCATTGACCCATCATTTGCTGTGTCAAG	798
Q	y	838	GACGGAAAGCATCACTCATGAAGGAAGAGATTTTGGCTGAAGACTGTGTACTCTCT	897
D	b	799	GACGGGAAGATGATCACTTACGAAGGAAGAGAGATTCGCTGGAAGAGCTTTGTACACC	858
Q	y	898	CCAGATCTGGTGTGCTTTTGTGGTGTAGAAATGTCCAGATGAAGCTTTCAATCAACCC	957
D	b	859	CCAGATCTTGGTCTTGATTCATCGTGGTAGAGTGTCTGATGAGGATTCATCTTGCCC	918
Q	y	958	ATCTGTGAGAATGCCACCTTTTACAGAGTACCAAGGAAGGCAGATGCCCCGTGGCTTG	1017
D	b	919	ATCTGTGAACGACACCTTTTAAAGGTACCAAGGCTGATGACCTGTGGCGCTG	978
Q	y	1018	GTGGTTCAATGGCCCCAGCATCTGTGTTGTGACAGCAGGTACACAGAGTGGATGGAG	1077
D	b	979	GTGGTCCATAGCCCCAGAATCTGTACTATCAGACAGCAGATACCAAGCAGTGGATGGAG	1038
Q	y	1078	AGGTTTGGGCTTGACACCCAGACACTTGTCTGTAATGAGACTGTGCCCTCAGTTTCAAC	1137
D	b	1039	AGGTTGGGCTTGACACAGCACTGATTTCTGAATGAGAATTGCCCTTCGGTCCACAC	1098
Q	y	1138	CTTCGAGGCCAAGAATTCAAAACCCAGCTCAACCTTCACCCGGGACATCTTCCCCCTG	1197
D	b	1099	CTGGCAGCCACBAGATTGACAGCCAGCTCAGCCTCATCCACCCTGACATCTTCCCCCAG	1158
Q	y	1198	CTCACAGTTTTCGTGTGAAGAGAGGGGCCACCTCTCAGTGTGCCCATGTTTCAAGGT	1257
D	b	1159	CTTACAGCTTCTATAGTAAGGAGAAAGGTTCCACCTTCAGCTGCAACAGATTCGGGGT	1218

US-09-564-805-221

Db	634	AATGGGC-----AGTCCCAACAGGAAGCATGGGCGAGGAC-CCTCCTTA	678
Qy	718	GTGCTAGCTTTTCATCTGTAAGCTTCACCTAAAGAGAGAAACTTCTTGGTGTCTCAAGCA	777
Db	679	GTGGTAGCTTTTGTCTGCAAGCTTCACCTGAGGAAGAAACTTCTTGGTGTCTTAAGCA	738
Qy	778	AAGGAGATGGGCCTCCACGTTGGGACAGCTGCCATCGCTCCCATCATTGTGCTGTCAAG	837
Db	739	AAGGAGCTGGGCCTTCTGTTGGGACGCGCCCATTGCACCCATCATTTGCTGTGTCAAG	798
Qy	838	GACGGGAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCT	897
Db	799	GACGGGAAGAGTATCACTTACGAAGGAAGAGAGATTGCTGCTGAAGAGCTTTGTACACCC	858
Qy	898	CCAGATCTCTGGTGTGCTTTTGTGTGTGTAGAAATGTCAGATGAAAGCTTCATTCAACCC	957
Db	859	CCAGATCTCTGGTCTGTATTTCATCTGTGTAGAGTGTCTGATGAAGGATTTCACTCTGCC	918
Qy	958	ATCTGTGAGAAAGCCACCTTTACAGAGGTACCAAGGAAGGCAGATGCCCGGTGGCCTTG	1017
Db	919	ATCTGTGAAGACGACACCTTTAAAGAGTACCGGACAGGCTGATGCACCTGTGGCGCTG	978
Qy	1018	GTGGTTTCAGATGGCCCCAGCATCTGTGCTGTGGACAGCAGGTACCAAGCAGTGGATGGAG	1077
Db	979	GTGGTCCATATAGCCCCAGAAATCTGTACTCATTCGACAGCAGATACCAAGCAGTGGATGGAG	1038
Qy	1078	AGGTTTGGGCCTGACACCCAGCAGCTTGTGCTCGAATGAGAAGTGTGCCTCAGTTTCACAAC	1137
Db	1039	AGGTTGGGCCTGACACAGCAGACCTGATTTCTGAATGAGAATTTGCCCTCTGGTCCACAAC	1098
Qy	1138	CTTCGAGCCACAAGATTCAAAACCGAGCTCAACCTCATCCACCGGACATCTTCCGCCCTG	1197
Db	1099	CTGCGAGCCACAAGATTCAAGACCCAGCTCAGCCTCATCCACCCCTGACATCTTCCCCCAG	1158
Qy	1198	CTCACAGTTTCCGTGTAAGGAAGAGGGCCCAACCTCTCAGTGTGCCATGGTTCAGGGT	1257
Db	1159	CTTACAGAGTTCTATAGTAAGGAGAAAGGTTCCACCTTCAGCTGCGCAACAGATTCGGGGT	1218

Qy	1078	AGGTTGGGCCCTTGACACACAGACACCTGATCTCTGAATGAGAACTGGCCCTGAGTCTGACCAAC	1115
Db	1039	AGGTTGGGGCCTTGACACACAGACACCTGATCTCTGAATGAGAACTGGCCCTCGGTCCCAAC	1098
Qy	1138	CTTCGAGGCCACAAGATTCAAACCCAGCTCAACCTCTATCCACCCGGACATCTTCCCCCTG	1197
Db	1099	CTGGGAGGCCACAGATTTCAGACCCAGCTCAGCCTCATCCACCTGACATCTTCCCCCG	1158
Qy	1198	CTACCAAGTTTCGGTGTGAAGAGAGGGGCCCAACCTTCAGTGTGCCATGGTTTCAGGT	1257
Db	1159	CTTACACAGTTCTTATAGTAAGGAGAAAGGGTCCACCTTCAGCTGTCCAAACAGTTTCGGGT	1218

[illegible]

Qy 1258 GAATGCTCTCTCAAGTACCAGTCCGTCCAGGAGGAGTGGCAGAGGATGCCATTATT 1317
 Db 1219 GAATGCTCTCTCAAGTATTCAAGTCCGCCCAAGAGAGAGTGGCAGAGGATACCAACATC 1278
 Qy 1318 ACTTGCATCTCTGAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCAATCTCCAGCAGAGC 1377
 Db 1279 GACTGCATCTGATGAATTCATAGTCAAGGCTTGGAGCTCCCAAGTTTCCAGGAGAT 1338
 Qy 1378 GTGAGGAGTACAGAGGAGTGGCAGAGAGCGGCCAGCCCCAGCAGCAGAGAGAAAAGAGTCAG 1437
 Db 1339 GTGAGGAGTATCGAAGAACCTGACGAGAAAACCCAGCCCCAGCAGAGAGAAAAGAGCCAG 1398
 Qy 1438 TACCAGAAATCATCTCTTGGACAGGCTCGCATCCCGATCCGATGAAGATTGCAATGTC 1497
 Db 1399 TATCTGAAATTTCTCTCGGTACGGGGTCTGCCATCCCAATGGAGATCCGAAATGTC 1458
 Qy 1498 AGTGCCACACTTGTCAACATAAGCCCCAGACAGCTCTCTGCTACTGCACTGTGGTGAGGGC 1557
 Db 1459 AGTTCCACACTGTCACCTAAGCCCTGACAAGTCAGTGTCTCTGATTTGGAGAGGC 1518
 Qy 1558 ACATTTGGCAGCTGTGCGCTCATTAACGAGACACAGGTGGACAGGCTCTGGGCAACCTG 1617
 Db 1519 ACTTTGGCAGTGTGCGCTCATTAACGAGACAGAAATAGACCGAGTCTTATGCAAGCCTC 1578
 Qy 1618 CTGCTGTGTTGTGTCCTCCACTGACAGCAGATCACACACGGGCTTCCCAAGTATCTTG 1677
 Db 1579 ACGGCTGTGTTGTGTCCTCCACTGACCGCCAGCACACAGGGCTTCTGAATATCTTG 1638
 Qy 1678 CTGACAGAGAAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTGTGCTGGTGT 1737
 Db 1639 CTGACAGAGAGCATGCGTTGGCATCTCTGGGAAACCTTCCAGCCCTTCTGTTGGTG 1698
 Qy 1738 GCCCCCAACAGCTCAAGCTGGCTCCAGCAGTACCAACCAAGTGCCAGAGGTCCTG 1797
 Db 1699 GCTCTACCCAGCTCAGGCGCTTGGCTGAGCATCTCTGGGAAACCTTCCAGCCCTTCTG 1758
 Qy 1798 CACCATCATGATGATTCCTGCCAAATGCTTCCAGGAGGCTGAGATCTCCAGTCT 1857
 Db 1759 CACCATCATGATGATTCCTGCCAAATGCTTCCAGGAGGCTGAGATCTCCAAATCT 1818
 Qy 1858 GCAGTGGAAAGATTCATGATGCTGCTTGGGAAACATGATGATTTGGAAGAGTTTCAGACC 1917
 Db 1819 ACATTTGGAAGGCTGATAGCTTGTGTTGGAAACATGATGATTTAGAGAAATTCAGACC 1878
 Qy 1918 TGTCTGGTGGGCACTGGAAGCATGCTTGGCTGTGGCTGGTGCAACCTCTGGCTGG 1977
 Db 1879 TGCCTGGTACGGCACTGCAAGCATGCTTTGGCTGTGCACTGGTACATTCATCTGGCTGG 1938
 Qy 1978 AAAGTGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGAAAGAT 2037
 Db 1939 AAAGTGTCTATTCCGGGGATACCATGCCCTGTGAGGCTCTGGTCCAGATGGGAAAGAT 1998
 Qy 2038 GCCACCTCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAAAGAGAGCAGTGAA 2097
 Db 1999 GCCACCTCTCTGATACATGAAGCCACTCTGAGGATCNCCTTGGAAAGAGAGCAGTAGAG 2058
 Qy 2098 AAGACACAGACAAAGCTGCCAAGCCATCAGCTGGGGATGCGATGAACCGGAGTTC 2157
 Db 2059 AGGACACACAGACAAAGCTGCCAAGCCATTAATGTTGGGATGCGGATGAATGCGGAGTTC 2118
 Qy 2158 ATTATGCTGAACCACTTCAGCAGGCTATGCCAGGTCCCTCTTCCAGCCCACTTC 2217
 Db 2119 ATCATGTGAACCACTTCAGTACGGGGTACGCAAGATTCCTCTTCCAGCCCTGACTTC 2178
 Qy 2218 AGCGAGAAAGTGGAGTTGCTTTCACACATGAAGGCTGCTTTGGAGACTTTTCCAAACA 2277
 Db 2179 AACGAGAAAGTTGGCATCGCTTTGACACATGAAGTCTGNTTTGGAGACTTCCGACA 2238
 Qy 2278 ATGCCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGCGGCACATCAGAGAGTAGAG 2337
 Db 2239 GTGCCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGAGGAGTGCATTTGAAGAGATGGTG 2298

Qy 2338 GAGCGAGGAGAGAGCGGAGCTCGCGAGTGGCGGCGCCCTCTCTGTCCAGGAGCTG 2397
 Db 2299 GAACGAGGAGAGAGAGGAGCTACGGTGGTGGAGAGCCCTCTCTGACC---CAGCAG 2355
 Qy 2398 GCAGGCGCTTGGAGATGGGAGCTTACAGAGCGGCGCCACACAGAGGAGCCACA 2456
 Db 2356 GCAGACGCCAGAGGACAGAGAACCCCAACAGAAAGCGGCGCCACACAGATGAACACCA 2414

RESULT 6
 US-09-564-805-210
 ; Sequence 210, Application US/09564805
 ; Patent No. 6333403
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.
 ; APPLICANT: Myriad Genetics, Inc.
 ; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
 ; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
 ; FILE REFERENCE: 2318-258
 ; CURRENT APPLICATION NUMBER: US/09/564,805
 ; CURRENT FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/107,468
 ; PRIOR FILING DATE: 1998-11-06
 ; PRIOR APPLICATION NUMBER: 09/434,382
 ; PRIOR FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 210
 ; LENGTH: 350
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (51)..(293)
 US-09-564-805-210

Query Match 10.0%; Score 247.4; DB 4; Length 350;
 Best Local Similarity 97.7%; Pred. No. 2.1e-57;
 Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGGGCGCTTCTCGCTGCTCGCGTCCGCGCGCGAGCAGCACCATGTTCGACGGAGCG 60
 Db 51 ATGTGGGCGCTTCTCGCTGCTCGCGTCCGCGCGCGAGCAGCACCATGTTCGACGGAGCG 110
 Qy 61 ACCATATCGCAGGACCAACCGCGCGCGAGCGCGCGCAAGGACCCGCTCGGCACCTG 120
 Db 111 ACCATATCGCAGGACCAACCGCGCGCGAGCGCGCGCAAGGACCCGCTCGGCACCTG 170
 Qy 121 CGCAGCGAGAGAGCGCGGACCGCTCGGGGTGCTCGGGCGGCCCAACACCGTGTACCTG 180
 Db 171 CGCAGCGAGAGAGCGCGGACCGCTCGGGGTGCTCGGGCGGCCCAACACCGTGTACCTG 230
 Qy 181 CAGGTGTGGCAGCGGTAGCCGGAGCTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
 Db 231 CAGGTGTGGCAGCGGTAGCCGGAGCTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 290
 Qy 241 AACCGGTATCTCTTCAA 257
 Db 291 AACCGGTATCTCTTCAA 307

RESULT 7
 US-09-564-805-28
 ; Sequence 28, Application US/09564805
 ; Patent No. 6333403
 ; GENERAL INFORMATION:
 ; APPLICANT: Tavtigian, Sean V.
 ; APPLICANT: Teng, David H.F.
 ; APPLICANT: Simard, Jacques
 ; APPLICANT: Rommens, Johanna M.

EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 315
LENGTH: 238
TYPE: DNA
ORGANISM: Homo sapiens
US-09-328-111-315

Query Match 9.6%; Score 237; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.le-54;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 ACCTGCAGGTGTGCACGGGTAGCCGGGACTGGGCGCGCGCTCTACGTCTCTCCG 235
Db 1 ACCTGCAGGTGTGCACGGGTAGCCGGGACTGGGCGCGCGCTCTACGTCTCTCCG 60
Qy 236 AGTTCAACCGGTATCTTCAACTCTGAGAGGGGTTTCAGAGCTCATGAGGAGCACA 295
Db 61 AGTTCAACCGGTATCTTCAACTCTGAGAGGGGTTTCAGAGCTCATGAGGAGCACA 120
Qy 296 AGTTAAAGTTGCTCGCTCGACACATATTCCTGCACAGCAATGCACTGCTTAATGTTG 355
Db 121 AGTTAAAGTTGCTCGCTCGACACATATTCCTGCACAGCAATGCACTGCTTAATGTTG 180
Qy 356 GGGCTTAAGTGAATGATCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTAC 412
Db 181 GGGCTTAAGTGAATGATCTTACTTTAAAGGAACCGGGCTTCCAAAGTGTGTAC 237

RESULT 10

US-09-564-805-27
Sequence 27, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/564,805
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRIOR APPLICATION NUMBER: 09/434,382
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 655
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(228)
OTHER INFORMATION: exon 24
NAME/KEY: polyA signal
LOCATION: (636)..(641)
US-09-564-805-27

Query Match 9.2%; Score 228; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 5.le-52;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2254 GTCTGCTTTGGAGCTTCCCAATGCCAGCTGATCCCCCACTGAAGCCCTGTTT 2313
Db 1 GTCTGCTTTGGAGCTTCCCAATGCCAGCTGATCCCCCACTGAAGCCCTGTTT 60
Qy 2314 GCTGCGCATCGAGGATGAGGAGCGCAGGAGAGCGGGAGCTCGGCGAGGTGCGG 2373

Db 61 GCTGCGCATCGAGGATGAGGAGCGCAGGAGAGCGGAGAGCTCGGCGAGGTGCGG 120
Qy 2374 CGGCGCTCTCTGCTCCAGGAGCTGGCAGGCGGCTTGAGGATGGGAGCCTCAGCAGAAG 2433
Db 121 CGGCGCTCTCTGCTCCAGGAGCTGGCAGGCGGCTTGAGGATGGGAGCCTCAGCAGAAG 180
Qy 2434 CGGCGCCACACAGAGGAGCCACAGGCGCAAGAGGTTCAGAGCCCAAGTGA 2481
Db 181 CGGCGCCACACAGAGGAGCCACAGGCGCAAGAGGTTCAGAGCCCAAGTGA 228

RESULT 11

US-09-564-805-26
Sequence 26, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
FILE REFERENCE: 2318-258
CURRENT APPLICATION NUMBER: US/09/564,805
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: US 60/107,468
PRIOR FILING DATE: 1998-11-06
PRIOR APPLICATION NUMBER: 09/434,382
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 240
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 145
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(145)
OTHER INFORMATION: exon 23
US-09-564-805-26

Query Match 5.8%; Score 145; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 6e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2109 CACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAATTATGCTGAA 2168
Db 1 CACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAATTATGCTGAA 60
Qy 2169 CCACCTCAGCCAGCGCTATGCCAAGTCCCGCTCTTACGCCCCCACTTCACGCGAAGT 2228
Db 61 CCACCTCAGCCAGCGCTATGCCAAGTCCCGCTCTTACGCCCCCACTTCACGCGAAGT 120
Qy 2229 GGGAGTTGCCCTTTCACCACTGAAG 2253
Db 121 GGGAGTTGCCCTTTCACCACTGAAG 145

RESULT 12

US-09-564-805-16
Sequence 16, Application US/09564805
Patent No. 6333403
GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.F.
APPLICANT: Simard, Jacques
APPLICANT: Rommens, Johanna M.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
FILE REFERENCE: 2318-258

; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: exon 13
US-09-564-805-16

Query Match 5.6%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.5e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1080 GTTTGGGCTGACACCCAGCACTTGGTCTGAATGAGAACTGTGCTCAGTTCAACAACCT 1139
Db 1 GTTTGGGCTGACACCCAGCACTTGGTCTGAATGAGAACTGTGCTCAGTTCAACAACCT 60
Qy 1140 TCGCAGCCACAGATTCAAAACCCAGCTCAACCTATCCACCCGGACATCTTCCCCCTGCT 1199
Db 61 TCGCAGCCACAGATTCAAAACCCAGCTCAACCTATCCACCCGGACATCTTCCCCCTGCT 120
Qy 1200 CACCAAGTTCCGCTGAAG 1218
Db 121 CACCAAGTTCCGCTGAAG 139

RESULT 13

US-09-564-805-20
; Sequence 20, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: exon 17
US-09-564-805-20

Query Match 5.6%; Score 139; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.5e-28;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1521 CCCCACAGCTCTGCTACTGCTGAGTGGTGGGACATTTGGGCAGCTGTGCGGTCA 1580
Db 1 CCCCACAGCTCTGCTACTGCTGAGTGGTGGGACATTTGGGCAGCTGTGCGGTCA 60

Qy 1581 TTACGGAGACCAGGTGGACAGGGTCCTGGGACACCTGCTGCTGTGTGTGTGCCACCT 1640
Db 61 TTACGGAGACCAGGTGGACAGGGTCCTGGGACACCTGCTGCTGTGTGTGTGCCACCT 120
Qy 1641 GCAGCGAGATCACCACACG 1659
Db 121 GCAGCGAGATCACCACACG 139

RESULT 14

US-09-564-805-24
; Sequence 24, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(121)
; OTHER INFORMATION: exon 21
US-09-564-805-24

Query Match 4.9%; Score 121; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1909 TTTGAGACCTGCTGTCGCGCACTGCAAGCATGCGTTTGGCTGCGCTGTGTGCACACC 1968
Db 1 TTTGAGACCTGCTGTCGCGCACTGCAAGCATGCGTTTGGCTGCGCTGTGTGCACACC 60

Qy 1969 TCTGCTGGAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCTCCGGATG 2028
Db 61 TCTGCTGGAAGTGGTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTCTCCGGATG 120
Qy 2029 G 2029
Db 121 G 121

RESULT 15

US-09-564-805-10
; Sequence 10, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05

```
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver..2.0
; SEQ ID NO 10
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(120)
; OTHER INFORMATION: exon 7
US-09-564-805-10

Query Match      4.8%; Score 120; DB 4; Length 120;
Best Local Similarity 100.0%; Pred.No. 3.le-23;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 560 GTGACAGAGGAGGGAAGCACCACCAACCATGGCAGAGTCCAGAAAGGCGCTCTCAGCAGGC 619
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1  GTGACAGAGGAGGGAAGCACCACCAACCATGGCAGAGTCCAGAAAGGCGCTCTCAGCAGGC 60

Qy 620 TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 679
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61  TCAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATG 120

Search completed: May 17, 2003, 16:40:30
Job time : 153.603 secs
```

Result No.	Score	Query Match	Length	DB	ID	Description	
1	2481	100.0	2481	9	US-09-988-626-1	Sequence 1, Appli	
2	2481	100.0	2481	9	US-09-988-687-1	Sequence 1, Appli	
3	2481	100.0	2958	9	US-09-988-626-3	Sequence 3, Appli	
4	2481	100.0	2958	9	US-09-988-687-3	Sequence 3, Appli	
5	2455.4	99.0	2908	9	US-09-988-626-223	Sequence 223, App	
6	2455.4	99.0	2908	9	US-09-988-687-223	Sequence 223, App	
7	2442.6	98.5	2892	9	US-09-988-626-225	Sequence 225, App	
8	2442.6	98.5	2892	9	US-09-988-687-225	Sequence 225, App	
9	1645.6	66.3	2470	9	US-09-988-626-221	Sequence 221, App	
10	1645.6	66.3	2470	9	US-09-988-687-221	Sequence 221, App	
11	734.8	29.6	793	10	US-09-833-381-2039	Sequence 2039, A	
12	470.4	19.0	536	10	US-09-833-381-2038	Sequence 2038, A	
13	432.8	17.4	554	9	US-09-918-995-8996	Sequence 8996, Ap	
14	247.4	10.0	350	9	US-09-988-626-210	Sequence 210, App	
15	247.4	10.0	350	9	US-09-988-687-210	Sequence 210, App	
16	247.4	10.0	26664	9	US-09-988-626-28	Sequence 28, Appl	
17	247.4	10.0	26664	9	US-09-988-687-28	Sequence 28, Appl	
18	245	9.9	295	9	US-09-988-626-4	Sequence 4, Appli	
19	245	9.9	295	9	US-09-988-687-4	Sequence 4, Appli	

Db 61 ACCATATCGAGGACCCGCGCGCGAGCGCGCGCAAGGACCCGCTGCGGCACCTG 120
Qy 121 CGACGCGAGAGACGGCGGACCGTTCGGGTGCTCGGGGGCCCAAAACACCGTGTACCTG 180
Db 121 CGACGCGAGAGACGGCGGACCGTTCGGGTGCTCGGGGGCCCAAAACACCGTGTACCTG 180
Qy 181 CAGGTGTGGCAGCGGTAGCCGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
Db 181 CAGGTGTGGCAGCGGTAGCCGGACTCGGGCGCGCGCTCTACGTCTTCTCCGAGTTC 240
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTCAGAGACTCATCAGGAGCACAAAGTTA 300
Db 241 AACCGGTATCTCTTCAACTGTGGAGAGCGTTTCAGAGACTCATCAGGAGCACAAAGTTA 300
Qy 301 AAGTGTGCTCGCTCGGACAAATATCTCTGACACGAAATGCACTGTCTAAATTTGGGGG 360
Db 301 AAGTGTGCTCGCTCGGACAAATATCTCTGACACGAAATGCACTGTCTAAATTTGGGGG 360
Qy 361 TTAAGTGAATGATCTTACTTTTAAAGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Db 361 TTAAGTGAATGATCTTACTTTTAAAGAAACCGGGCTTCCAAAGTGTACTTTCTGGA 420
Qy 421 COTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Db 421 COTCCAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTTGAAAGGA 480
Qy 481 ATAGAACTGGTGTGCGGCCCACTCTGTCGCCAGAAATACGAGGATGAAACCATGACGTT 540
Db 481 ATAGAACTGGTGTGCGGCCCACTCTGTCGCCAGAAATACGAGGATGAAACCATGACGTT 540
Qy 541 TACAGATCCCATACAGTGAACAGAGGAGGGAAGCAACCACTGCGAGTGCA 600
Db 541 TACAGATCCCATACAGTGAACAGAGGAGGGAAGCAACCACTGCGAGTGCA 600
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGACTCCGAGTCGAAATGAAAT 660
Db 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTCAGACTCCGAGTCGAAATGAAAT 660
Qy 661 GAGCCACACTTCCACATGTTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGTC 720
Db 661 GAGCCACACTTCCACATGTTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGTC 720
Qy 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Db 721 GTAGCTTTTCACTGTAAGCTTCACTTAAAGAGAGAACTCTTGGTGTCAAAGCAAG 780
Qy 781 GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAAGAC 840
Db 781 GAGATGGGCTCCCACTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAAGAC 840
Qy 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 841 GGGAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 901 GATCTGTGTGCTTTTGTGTGTAGAACTCCAGATGAAGCTTCATTCACCCATC 960
Db 901 GATCTGTGTGCTTTTGTGTGTAGAACTCCAGATGAAGCTTCATTCACCCATC 960
Qy 961 TGTGAGAAATGCCACTTTTCAAGGTACCAAGAAAGGAGAGATGCCCGTGGCTTTGGT 1020
Db 961 TGTGAGAAATGCCACTTTTCAAGGTACCAAGAAAGGAGAGATGCCCGTGGCTTTGGT 1020
Qy 1021 GTTTCATGTGGCCCAAGCATCTGTGTGTGTGGACAGAGGTACAGAGTGTGATGAGAG 1080
Db 1021 GTTTCATGTGGCCCAAGCATCTGTGTGTGTGGACAGAGGTACAGAGTGTGATGAGAG 1080
Qy 1081 TTTGGGCTGACACCCAGCACTTGTCTCTGAATGAGAACTGTGCTCAGTTTCACAACTT 1140
Db 1081 TTTGGGCTGACACCCAGCACTTGTCTCTGAATGAGAACTGTGCTCAGTTTCACAACTT 1140
Qy 1141 CGAGCCCAAGATTTAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTGTC 1200
Db 1141 CGAGCCCAAGATTTAAACCCAGCTCAACCTCATCCACCCGAGCATCTTCCCTGTC 1200

Qy 1201 ACCAGTTTCGCTGTAAAGAGGGGCCCAACCTCAGTGTGCCCATGGTTTCAGGTTGAA 1260
Db 1201 ACCAGTTTCGCTGTAAAGAGGGGCCCAACCTCAGTGTGCCCATGGTTTCAGGTTGAA 1260
Qy 1261 TGCCTCTCTCAAGTACCAGTCCGTCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCTCTCTCAAGTACCAGTCCGTCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGTGCAGCTTCCAACTTCCAGCAGAGCGT 1380
Db 1321 TGCATCTCTGAGGAATTCATAGTTGAGGCGTGCAGCTTCCAACTTCCAGCAGAGCGT 1380
Qy 1381 CAGAGTACAGGAGGAGTGGCAGGACGGCCAGCCAGCCAGCAGAGAGAAAGTCAGTAC 1440
Db 1381 CAGAGTACAGGAGGAGTGGCAGGACGGCCAGCCAGCCAGCAGAGAGAAAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTTGGAACAGGCTTCCCATCCGATGAAAGATTCGAAATGTCACT 1500
Db 1441 CCAGAAATCATCTTCTTTGGAACAGGCTTCCCATCCGATGAAAGATTCGAAATGTCACT 1500
Qy 1501 GCCACACTTGTCAACATAAGCCCGCACACGCTCTCTGTACTTGGACTGTGTGAGGGCACA 1560
Db 1501 GCCACACTTGTCAACATAAGCCCGCACACGCTCTCTGTACTTGGACTGTGTGAGGGCACA 1560
Qy 1561 TTTGGGAGCTGTGCGCTCATTTACGGAGACAGGTGGACAGGGTCTCTGGGCACTTGGCT 1620
Db 1561 TTTGGGAGCTGTGCGCTCATTTACGGAGACAGGTGGACAGGGTCTCTGGGCACTTGGCT 1620
Qy 1621 GCTGTGTTTGTGTCCTTGGCAACAGGCTTCCACACAGGCTTCCCAAGTATCTTGTCTG 1680
Db 1621 GCTGTGTTTGTGTCCTTGGCAACAGGCTTCCACACAGGCTTCCCAAGTATCTTGTCTG 1680
Qy 1681 CAGAGAGAACGCGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTTTGTGTTGTTGCC 1740
Db 1681 CAGAGAGAACGCGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTTTGTGTTGTTGCC 1740
Qy 1741 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGGAGTCTCTGCAC 1800
Db 1741 CCCAACAGCTCAAAAGCTGTGCTCCAGCAGTACCAACACAGTGCAGGAGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTTCTGTCGCAATGCTTCCAGAGGGCTGAGATCTCCAGTCTTGA 1860
Db 1801 CACATCAGTATGATTTCTGTCGCAATGCTTCCAGAGGGCTGAGATCTCCAGTCTTGA 1860
Qy 1861 GTGAAAGATTTGATCAGTTTCTGTCGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Db 1861 GTGAAAGATTTGATCAGTTTCTGTCGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Qy 1921 CTGCTGGGCACCTCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACTCTGGCTGGA 1980
Db 1921 CTGCTGGGCACCTCAAGCATGCGTTTGGCTGTGCGCTGTGTCACACTCTGGCTGGA 1980
Qy 1981 GTGCTCTATTCCGGGGACACCATCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGCTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGGTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGTATACATGAAGCCACCTCTGGAAGATGTTTGGAGAGGAGCAAGTGGAAAG 2100
Db 2041 ACCCTCTGTATACATGAAGCCACCTCTGGAAGATGTTTGGAGAGGAGCAAGTGGAAAG 2100
Qy 2101 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGGGAGTTCAAT 2160
Db 2101 ACACAGCACAAGCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTTCAGCCCCCACTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCTCTTTCAGCCCCCACTCAGC 2220
Qy 2221 GAGAAAGTGGAGTGTCTTTGACCAATGAAGTCTGTCTTTGAGACTTTTCCAAACATG 2280
Db 2221 GAGAAAGTGGAGTGTCTTTGACCAATGAAGTCTGTCTTTGAGACTTTTCCAAACATG 2280

QY 2281 CCCAAGCTGATTCCCTCCCACTGAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340
DB 2281 CCCAAGCTGATTCCCTCCCACTGAAGCCCTGTTGCTGGGACATCGAGAGATGGAGGAG 2340
QY 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTCTGTCACAGGAGCTGGCA 2400
DB 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGCGGCCCTCTGTCACAGGAGCTGGCA 2400
QY 2401 GGGGGCTCGAGATGGGAGCTTACGACAGAGCGGGCCCAACACAGAGAGCCACAGGCC 2460
DB 2401 GGGGGCTCGAGATGGGAGCTTACGACAGAGCGGGCCCAACACAGAGAGCCACAGGCC 2460
QY 2461 AAGAAGTCCAGAGCCCAAGTGA 2481
DB 2461 AAGAAGTCCAGAGCCCAAGTGA 2481

RESULT 2

US-09-988-687-1
; Sequence 1, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-1

Query Match 100.0%; Score 2481; DB 9; Length 2481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGGCGCTTTGCTCGCTGCTGGGTCCGGCCGGGCGGACGACCATGTCGAGGAGCC 60
DB 1 ATGTGGCGCTTTGCTCGCTGCTGGGTCCGGCCGGGCGGACGACCATGTCGAGGAGCC 60
QY 61 ACCATATCGAGGACCCCGCCCGCGAGCGGCGCGCAAGACCGCTGCGGACCTG 120
DB 61 ACCATATCGAGGACCCCGCCCGCGAGCGGCGCGCAAGACCCGCTGCGGACCTG 120
QY 121 CGCACGCGAGAGAAGCGGACCGTTCGGGGTGTCCGGCGGCCCAACACCGGTGTACCTG 180
DB 121 CGCACGCGAGAGAAGCGGACCGTTCGGGGTGTCCGGCGGCCCAACACCGGTGTACCTG 180
QY 181 CAGGTGGTGGACGGGTAGCGGAGTCTCGGGCGCGCGCTCTAGCTTCTCCGAGTTC 240
DB 181 CAGGTGGTGGACGGGTAGCGGAGTCTCGGGCGCGCGCTCTAGCTTCTCCGAGTTC 240
QY 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATCGAGGACCAAGTTA 300
DB 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATCGAGGACCAAGTTA 300

QY 301 AAGTTGCTCGCTCGGACAAATATTTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 360
DB 301 AAGTTGCTCGCTCGGACAAATATTTCTTGACACGAATGCACTGGTCTTAATGTTGGGGC 360
QY 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGCTACTTTCTGA 420
DB 361 TTAAGTGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGCTACTTTCTGA 420
QY 421 CTTCCACAATCGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
DB 421 CTTCCACAATCGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
QY 481 ATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGAGGATGAACCAATGACAGTT 540
DB 481 ATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGAGGATGAACCAATGACAGTT 540
QY 541 TACCAGATCCCCATACAGTGAACAGAGGAGGAAAGCACCAACCATGGCAGAGTCCA 600
DB 541 TACCAGATCCCCATACAGTGAACAGAGGAGGAAAGCACCAACCATGGCAGAGTCCA 600
QY 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 660
DB 601 GAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCCGAATGAAAT 660
QY 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
DB 661 GAGCCACACTTCCACATGGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
QY 721 GTAGCTTTCTCTGTAAGCTTCACTTTAAGAGAGAACTTCTTGGTCTCAAGCAAG 780
DB 721 GTAGCTTTCTCTGTAAGCTTCACTTTAAGAGAGAACTTCTTGGTCTCAAGCAAG 780
QY 781 GAGATGGGCTCCCACTGGGACACCTGCTCCATCGCTCCCATCATTTGCTGCTCTCAAGGAC 840
DB 781 GAGATGGGCTCCCACTGGGACACCTGCTCCATCGCTCCCATCATTTGCTGCTCTCAAGGAC 840
QY 841 GGGAAAGCATCACTCATGAAGAAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCA 900
DB 841 GGGAAAGCATCACTCATGAAGAAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCA 900
QY 901 GATCCTGGTGTGCTTTTGGTGTGTAATGTCAGATGAAGCTTCAATCAACCCATC 960
DB 901 GATCCTGGTGTGCTTTTGGTGTGTAATGTCAGATGAAGCTTCAATCAACCCATC 960
QY 961 TGTGAGAAATGCCACCTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCGCTGGCTTGGTG 1020
DB 961 TGTGAGAAATGCCACCTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCGCTGGCTTGGTG 1020
QY 1021 GTTCACATGGCCCCAGCATCTGTGTTGTGACACAGGTACCCAGAGTGGATGGAGAGG 1080
DB 1021 GTTCACATGGCCCCAGCATCTGTGTTGTGACACAGGTACCCAGAGTGGATGGAGAGG 1080
QY 1081 TTTGGGCTGACACCCAGCATTTGCTCTGATGAGAACTGTGCTCAGTTCAGTTCAACCTT 1140
DB 1081 TTTGGGCTGACACCCAGCATTTGCTCTGATGAGAACTGTGCTCAGTTCAGTTCAACCTT 1140
QY 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCACTCCACCGGACATCTTCCCCCTGCTC 1200
DB 1141 CGCAGCCAAAGATTCAAAACCCAGCTCAACCTCACTCCACCGGACATCTTCCCCCTGCTC 1200
QY 1201 ACCAGTTTCCGCTGTAAGAAAGGAGGCCCCACCTCAGTGTGCCCATGGTTCAGGGTGA 1260
DB 1201 ACCAGTTTCCGCTGTAAGAAAGGAGGCCCCACCTCAGTGTGCCCATGGTTCAGGGTGA 1260
QY 1261 TGCCTTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTTATTACT 1320
DB 1261 TGCCTTCTCAAGTACAGCTCCGTCAGGAGGAGTGGCAGAGGATGCCATTTATTACT 1320
QY 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCACTTCAGCAGAGCGTG 1380
DB 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGCAAGCTTCCCACTTCAGCAGAGCGTG 1380
QY 1381 CAGGAGTACAGGAGGAGTCCGACGAGCGGCCCGCAGCCCGCAGCAGAGAGAAAGAGTAC 1440

Db 1381 CAGGAGTACAGGAGGTGCGCAGACGGCCAGCCCGCAGCAGAGAAAGAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCCCTTGGAAACAGGGCTGCGCATCCCGATGAAGATTGCAAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCCCTTGGAAACAGGGCTGCGCATCCCGATGAAGATTGCAAAATGTCAGT 1500
Qy 1501 GCCACACTGTCAACATGAAGCCCGACAGCTCTGCTACTGGACTGTGGTGGGGCACA 1560
Db 1501 GCCACACTGTCAACATGAAGCCCGACAGCTCTGCTACTGGACTGTGGTGGGGCACA 1560
Qy 1561 TTTGGGCGAGCTGTGCGGTCATTACCGAGACACAGGTGGACAGGGTCTCGGCACCCCTGGCT 1620
Db 1561 TTTGGGCGAGCTGTGCGGTCATTACCGAGACACAGGTGGACAGGGTCTCGGCACCCCTGGCT 1620
Qy 1621 GCTGTGTTTGTGTCCTCCACTGCACGACATCACCAACCGGGTTCGCCAAGTATCTTGTCTG 1680
Db 1621 GCTGTGTTTGTGTCCTCCACTGCACGACATCACCAACCGGGTTCGCCAAGTATCTTGTCTG 1680
Qy 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGAAAGCCGCTTACCCCTTTGCTGGTGGTTGCC 1740
Db 1681 CAGAGAGAACGGCCCTTGGCATCTTTTGGAAAGCCGCTTACCCCTTTGCTGGTGGTTGCC 1740
Qy 1741 CCCAACCAAGCTCAAAAGCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGTCTCTGCAC 1800
Db 1741 CCCAACCAAGCTCAAAAGCTGGCTCCAGCAGTACCAACACCAAGTGCAGAGTCTCTGCAC 1800
Qy 1801 CACATCAGTATGATTCCTGCCAAATSCCTTCAGGAAGGGGTGAGATCTCCAGTCCCTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCCAAATSCCTTCAGGAAGGGGTGAGATCTCCAGTCCCTGCA 1860
Qy 1861 GTGAAAGATTGATCAGTTCGCTGTGCGAATGTGTAATTTGGAAGAGTTTCAGACCTGT 1920
Db 1861 GTGAAAGATTGATCAGTTCGCTGTGCGAATGTGTAATTTGGAAGAGTTTCAGACCTGT 1920
Qy 1921 CTGTTGGCGGCTCAAGCATGCTGCTTGGCTGTGCGTGTGTCACACCTCTGGCTGGAAA 1980
Db 1921 CTGTTGGCGGCTCAAGCATGCTGCTTGGCTGTGCGTGTGTCACACCTCTGGCTGGAAA 1980
Qy 1981 GTGTTCTATTCCGGGACACCATGCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
Db 1981 GTGTTCTATTCCGGGACACCATGCTGCGAGGCTCTGGTCCGGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGTGGTTTGGAGAGAGCAGTGGAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGGAAAGTGGTTTGGAGAGAGCAGTGGAAG 2100
Qy 2101 ACACACAGCACAAGCTGCCAAGCCATCAGCGTGGGATGCGGATGAACCGGAGTTCAAT 2160
Db 2101 ACACACAGCACAAGCTGCCAAGCCATCAGCGTGGGATGCGGATGAACCGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGSTCCCCCTCTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGCTATGCCAAGSTCCCCCTCTTCAGCCCCCACTTCAGC 2220
Qy 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Db 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Qy 2281 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGCGACATCAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTCCTCCCACTGAAGCCCTGTTTGTGCGGACATCAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGGGGCCCTCTTCAGGAGAGCTGGCA 2400
Db 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGCGGGGGGCCCTCTTCAGGAGAGCTGGCA 2400
Qy 2401 GCGGCTCTGGAGATGGGAGCTCAGCAGAGAGCGGGGCCACACAGAGAGGCCACAGGCC 2460
Db 2401 GCGGCTCTGGAGATGGGAGCTCAGCAGAGAGCGGGGCCACACAGAGAGGCCACAGGCC 2460
Qy 2461 AAGAAAGTTCAGAGCCAGTGA 2481

Db 2461 AAGAAAGTTCAGAGCCAGTGA 2481
RESULT 3
US-09-988-626-3
; Sequence 3, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51)..(2531)
; OTHER INFORMATION: coding sequence as in SEQ ID NO:1
US-09-988-626-3

Query Match 100.0%; Score 2481; DB 9; Length 2958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGGGCGCTTTGCTCGCTGCTCGGTCCGGCGCGCGGACGCCACCATGTCGACGGGAGCG 60
Db 51 ATGTGGGCGCTTTGCTCGCTGCTCGGTCCGGCGCGGACGCCACCATGTCGACGGGAGCG 110
Qy 61 ACCATATCGAGGACCCCGCCCGCGAGCGCGCGGCGGCAAGGACCCGCTGCGGCACTG 120
Db 111 ACCATATCGAGGACCCCGCCCGCGAGCGCGCGGCGGCAAGGACCCGCTGCGGCACTG 170
Qy 121 CGCACGCGAGAGCGCGGACCGTCCGGGGTGTCCGGCGGCCCAACACCGCTGTACCTG 180
Db 171 CGCACGCGAGAGCGCGGACCGTCCGGGGTGTCCGGCGGCCCAACACCGCTGTACCTG 230
Qy 181 CAGGTGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGCTCTTCCGAGTTC 240
Db 231 CAGGTGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGCTCTTCCGAGTTC 290
Qy 241 AACCGGTATCTCTTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGACACAAGTTA 300
Db 291 AACCGGTATCTCTTCAACTGTGGAGAGGGCTTCAGAGACTCATGCGAGGACACAAGTTA 350
Qy 301 AAGGTTGCTCGCTGGCAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
Db 351 AAGGTTGCTCGCTGGCAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 410
Qy 361 TTAAGTGAATGATCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 420
Db 411 TTAAGTGAATGATCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGA 470
Qy 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 480
Db 471 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGA 530
Qy 481 ATGAACTGGGCTGTGCGGCCCCACTCTGCCCCAGAAATACGAGGATGAACCATGACACTT 540

Db 531 ATAGAACTGGCTGTGCGGCCCACTCTGCCCGAGAAATACGAGATGAACCAATGACAGTT 590
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCAATGGCAGAGTCCA 600
Db 591 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCAATGGCAGAGTCCA 650
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTACAGTCCGAGTCGAATAAAT 660
Db 651 GAAAGGCTCTCAGCAGGCTCAGTCAGAGCGATCTTACAGTCCGAGTCGAATAAAT 710
Qy 661 GAGCACACCTTCCACATCGTGTAGCAGAGAGAGGGGTGAGGACCTTCCCTGTGC 720
Db 711 GAGCACACCTTCCACATCGTGTAGCAGAGAGAGGGGTGAGGACCTTCCCTGTGC 770
Qy 721 GTAGCTTTTCATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCAAAGCAAG 780
Db 771 GTAGCTTTTCATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCAAAGCAAG 830
Qy 781 GAGATGGGCTCCAGTTGGGACAGTCGCCATCGCTCCCATATTGCTGCTCAGAGAC 840
Db 831 GAGATGGGCTCCAGTTGGGACAGTCGCCATCGCTCCCATATTGCTGCTCAGAGAC 890
Qy 841 GGGAAAAGCATCACTCATCAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 891 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 950
Qy 901 GATCTGTGTGCTCTTTTGTGTGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCATC 960
Db 951 GATCTGTGTGCTCTTTTGTGTGTGTAGATGTCCAGATGAAAGCTTCAATCAACCCATC 1010
Qy 961 TGTGAAATGCGACCTTTTCAAGAGTACCAAGAAAGGAGATGCCCTGCTGCTGTGGTG 1020
Db 1011 TGTGAAATGCGACCTTTTCAAGAGTACCAAGAAAGGAGATGCCCTGCTGCTGTGGTG 1070
Qy 1021 GTTCATGGCCCGACATCTGTGTGTGTGTGAGAGAGGATACAGAGTGGATGGAGG 1080
Db 1071 GTTCATGGCCCGACATCTGTGTGTGTGTGAGAGAGGATACAGAGTGGATGGAGG 1130
Qy 1081 TTTGGGCTTGACACCCAGACATCTGTGTGTGTGTGAGATGAACTGTGCTCAGTTCACAACTT 1140
Db 1131 TTTGGGCTTGACACCCAGACATCTGTGTGTGTGTGAGATGAACTGTGCTCAGTTCACAACTT 1190
Qy 1141 CGAGCCACAAGATTCAAACCCAGTCAACCTCAATCCACCGGACATCTTCCCTGCTC 1200
Db 1191 CGAGCCACAAGATTCAAACCCAGTCAACCTCAATCCACCGGACATCTTCCCTGCTC 1250
Qy 1201 ACCAGTTCCGCTGTAAAGAGAGGGCCCAACCTCAGTGTGCCATGTTTCAAGGTGA 1260
Db 1251 ACCAGTTCCGCTGTAAAGAGAGGGCCCAACCTCAGTGTGCCATGTTTCAAGGTGA 1310
Qy 1261 TGCTCTCAAGTACCAAGTCCGCTCCAGAGGAGTGGCAGAGGATGCCATTATTACT 1320
Db 1311 TGCTCTCAAGTACCAAGTCCGCTCCAGAGGAGTGGCAGAGGATGCCATTATTACT 1370
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCAGCAGAGCGTG 1380
Db 1371 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGACGCTTCCCACTTCCAGCAGAGCGTG 1430
Qy 1381 CAGGAGTACAGAGGAGTGGCAGAGGCGGCCCGCCAGCCGACGAGAAAGATGATGATAC 1440
Db 1431 CAGGAGTACAGAGGAGTGGCAGAGGCGGCCCGCCAGCCGACGAGAAAGATGATGATAC 1490
Qy 1441 CCAGAAATCATCTTCTTGGAAAGAGGCTGTGCCATCCGATCCGATTCGAAATGTCACT 1500
Db 1491 CCAGAAATCATCTTCTTGGAAAGAGGCTGTGCCATCCGATCCGATTCGAAATGTCACT 1550
Qy 1501 GCCACACTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGACCTGTGGTGAGGCGACA 1560
Db 1551 GCCACACTTGTCAACATAAGCCCGACAGCTCTCTGCTACTGACCTGTGGTGAGGCGACA 1610
Qy 1561 TTTGGCAGCTGTGCGCTCATTTACGAGACACAGGTGGAAGGCTCTGGGACCTGCTGGCT 1620

Db 1611 TTTGGCAGCTGTGCGCTCATTTACGAGACACAGGTGGACAGGGTCTTGGGACACCTGGCT 1670
Qy 1621 GCTGTGTTTGTGTCCACCTGCACGAGATCACACACGGGCTTCCCAAGTATCTTGTG 1680
Db 1671 GCTGTGTTTGTGTCCACCTGCACGAGATCACACACGGGCTTCCCAAGTATCTTGTG 1730
Qy 1681 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTCTGCTGCTGGTGGC 1740
Db 1731 CAGAGAGAACGGCTTGGCATCTTTGGGAAAGCGCTTACCCCTTCTGCTGCTGGTGGC 1790
Qy 1741 CCCAACACAGCTCAAAAGCTGCTCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1800
Db 1791 CCCAACACAGCTCAAAAGCTGCTCCAGCAGTACCAACACAGTGCAGAGGTCTCTGCAC 1850
Qy 1801 CACATCAGTATGATTCCTGCGCAAAATGCCCTTCAAGAAAGGGCTGAGATCTCCAGTCTGCA 1860
Db 1851 CACATCAGTATGATTCCTGCGCAAAATGCCCTTCAAGAAAGGGCTGAGATCTCCAGTCTGCA 1910
Qy 1861 GTGGAAAGATTGATCAGTTCGCTGTTCGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1920
Db 1911 GTGGAAAGATTGATCAGTTCGCTGTTCGCAACATGTGATTTGGAAAGATTTTCAGACCTGT 1970
Qy 1921 CTGTGCGGCACTGCAAGCATGCTTTTGGCTGTGCTGTGGTGCACACCTCTCGCTGGAAA 1980
Db 1971 CTGTGCGGCACTGCAAGCATGCTTTTGGCTGTGCTGTGGTGCACACCTCTCGCTGGAAA 2030
Qy 1981 GTGTCTATTCCGGGACACCATGCCCTGCGAGGCTCTGCTCGGATCGGGAAGATGCC 2040
Db 2031 GTGTCTATTCCGGGACACCATGCCCTGCGAGGCTCTGCTCGGATCGGGAAGATGCC 2090
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGAAGATGTTTGGAAAGAGAGACAGTGGAAAAG 2100
Db 2091 ACCCTCTGATACATGAAGCCACCTTGAAGATGTTTGGAAAGAGAGACAGTGGAAAAG 2150
Qy 2101 ACACACACACAACTGCTCCAAAGCCATCAGCGTGGGGATGCGGATGAAACGCGAGTTCA 2160
Db 2151 ACACACACACAACTGCTCCAAAGCCATCAGCGTGGGGATGCGGATGAAACGCGAGTTCA 2210
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTTATCCAAAGTTCCTTTCAGGCCCACTTCAGC 2220
Db 2211 ATGCTGAACCACTTTCAGCCAGCGCTTATCCAAAGTTCCTTTCAGGCCCACTTCAGC 2270
Qy 2221 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2280
Db 2271 GAGAAAGTGGAGTTGCTTTGACACATGAAGTCTGCTTTGGAGACTTTTCCAAATG 2330
Qy 2281 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2340
Db 2331 CCCAAGCTGATTTCCCACTGAAAGCCCTGTTGCTGGGACATCGAGAGATGAGGAG 2390
Qy 2341 CGCAGGAGAAAGCGGAGCTGCGGAGGTGGGGCGGCCCTCTCTCCAGGAGCTGGCA 2400
Db 2391 CGCAGGAGAAAGCGGAGCTGCGGAGGTGGGGCGGCCCTCTCTCCAGGAGCTGGCA 2450
Qy 2401 GCGGCTCTGGAGATGGGAGCTTCAAGAGAGCGGGGCCACACAGAGAGCCACAGGCC 2460
Db 2451 GCGGCTCTGGAGATGGGAGCTTCAAGAGAGCGGGGCCACACAGAGAGCCACAGGCC 2510
Qy 2461 AAGAGGTTCAGAGCCAGTGA 2481
Db 2511 AAGAGGTTCAGAGCCAGTGA 2531

RESULT 4

US-09-988-687-3
; Sequence 3, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigan, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.


```
Qy 901 GATCCTGTGCTGCTTTTGTGTGTAGAAATGTCAGATGAAAGCTTCATTCAACCCATC 960
Db 901 GATCCTGTGCTGCTTTTGTGTGTAGAAATGTCAGATGAAAGCTTCATTCAACCCATC 960
Qy 961 TGTGAGAAATGCCACTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCCCGTGGCTTGGTG 1020
Db 961 TGTGAGAAATGCCACTTTTCAGAGGTACCAAGAAAGGAGAGATGCCCCCGTGGCTTGGTG 1020
Qy 1021 GTTTCATGTGGCCCGCAGCATCTGTGCTTGTGGACACAGGTACACAGTGGATGGAGAGG 1080
Db 1021 GTTTCATGTGGCCCGCAGCATCTGTGCTTGTGGACACAGGTACACAGTGGATGGAGAGG 1080
Qy 1081 TTTTGGCCCTGACACCCAGCACTTGTGCTGTGAATGAGAACTGTGCCCTCAGTTTCAACAACCTT 1140
Db 1081 TTTTGGCCCTGACACCCAGCACTTGTGCTGTGAATGAGAACTGTGCCCTCAGTTTCAACAACCTT 1140
Qy 1141 CGCAGCCCAAGATTCAAAACCCAGCTCAACCTCATCCACCCGGAACATCTTCCCCCTGCTC 1200
Db 1141 CGCAGCCCAAGATTCAAAACCCAGCTCAACCTCATCCACCCGGAACATCTTCCCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAAGAGAGAGGCCCCCAGCCCTCAGTGTGCCATGGTTCAAGGGTAA 1260
Db 1201 ACCAGTTTCCGCTGTAAAGAGAGAGGCCCCCAGCCCTCAGTGTGCCATGGTTCAAGGGTAA 1260
Qy 1261 TGCCCTCTCAAGTACAGCTCCGCTCCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Db 1261 TGCCCTCTCAAGTACAGCTCCGCTCCAGAGGAGGTGGCAGAGGATGCCATTATTACT 1320
Qy 1321 TGCATCTCTGAGAAATTCATAGTTGAGCGCTGCGAGCTTCCAACTTCACAGAGCGGTG 1380
Db 1321 TGCATCTCTGAGAAATTCATAGTTGAGCGCTGCGAGCTTCCAACTTCACAGAGCGGTG 1380
Qy 1381 CAGGAGTACAGAGAGGTGCGCAGAGCGGCCAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1440
Db 1381 CAGGAGTACAGAGAGGTGCGCAGAGCGGCCAGCCCGCAGCAGAGAGAAAGAGTCAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCGATCCCGATTCGAAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTGGAAACAGGCTCTGCCATCCCGATTCGAAATGTCAGT 1500
Qy 1501 GCCACACTTGTCAACATAAGCCCGACACGCTCTCTGCTACTGGAATGTTGGTGAAGGCA 1560
Db 1501 GCCACACTTGTCAACATAAGCCCGACACGCTCTCTGCTACTGGAATGTTGGTGAAGGCA 1560
Qy 1561 TTTTGGCAGCTGTGCGCTCATTTACGAGACACAGGTGGACAGGCTCTGGGCACCTGGCT 1620
Db 1561 TTTTGGCAGCTGTGCGCTCATTTACGAGACACAGGTGGACAGGCTCTGGGCACCTGGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACCTGCACGAGATCACACACGGGCTTGCAGATCTTGTGCTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACCTGCACGAGATCACACACGGGCTTGCAGATCTTGTGCTG 1680
Qy 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCCGCTTCAACCTTGTGCTGTTGCC 1740
Db 1681 CAGAGAGAACCGCCTTGGCATCTTTTGGAAAGCCGCTTCAACCTTGTGCTGTTGCC 1740
Qy 1741 CCCAACCCAGCTCAAAAGCCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGTCTGGAC 1800
Db 1741 CCCAACCCAGCTCAAAAGCCTGGCTCCAGCAGTACCAACACAGTGGCAGAGGTCTGGAC 1800
Qy 1801 CACATCAGTATGATTTCTCCCAAATGCCCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1860
Db 1801 CACATCAGTATGATTTCTCCCAAATGCCCTTTCAGGAAGGGCTGAGATCTCCAGTCTCTGCA 1860
Qy 1861 GTGGAAGATTGATCAGTTTGTGCTGTGGAAACATGTGATTTTGGAAAGTTTCAGACCTGT 1920
Db 1861 GTGGAAGATTGATCAGTTTGTGCTGTGGAAACATGTGATTTTGGAAAGTTTCAGACCTGT 1920
Qy 1921 CTGGTGGCGCACTGAAGCATCGTTTGGCTGTGGCTGGTGCACACCTCTCGCTGGGAA 1980
Db 1921 CTGGTGGCGCACTGAAGCATCGTTTGGCTGTGGCTGGTGCACACCTCTCGCTGGGAA 1980
Qy 1981 GTGGTCTATTCCGGGGACACCATGCCCCCTCGAGGCTCTGTGTCGGATGGGAAAGATGCC 2040
```

```
Db 1981 GTGCTCTATTCCGGGGACACCATGCCCTCGAGGCTCTGTGTCGGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTCTGATACATGAAGCCACCCTCGAAGATGTTTGGNAGAGGAGCAAGTGGAAAG 2100
Db 2041 ACCCTCTCTGATACATGAAGCCACCCTCGAAGACGGTTTGGNAGAGGAGCAAGTGGAAAG 2100
Qy 2101 ACACACAGCAACGTCCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCATT 2160
Db 2101 ACACACAGCAACGTCCTCCCAAGCCATCAGCGTGGGATGCGGATGAACGCGGAGTTCATT 2160
Qy 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Qy 2221 GAGAAAGTGGAGTGTGCTTTGACCATGAAGTCTGCTTTGAGAGCTTTTGCACCAATG 2280
Db 2221 GAGAAAGTGGAGTGTGCTTTGACCATGAAGTCTGCTTTGAGAGCTTTTGCACCAATG 2280
Qy 2281 CCCAAGCTGATTTCCCCACTGAAAGCCTGTTTCTGCGACATCGAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCCACTGAAAGCCTGTTTCTGCGACATCGAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAGAGCGGAGCTCGCGCAGGTGCGCGCGCCCTCTCTGTCCAGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGAGCGGAGCTCGCGCAGGTGCGCGCGCCCTCTCTGTCCAGGAGCTGGCA 2400
Qy 2401 GCGCGCTTGGAGTGGGAGCCTCAGCAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Db 2401 GCGCGCTTGGAGTGGGAGCCTCAGCAGAGCGGCGCCACACAGAGGAGCCACAGGCC 2460
Qy 2461 AAGAAGTTCAGAGCCCACTGA 2481
Db 2461 AAGAAGTTCAGAGCCCACTGA 2481
```

RESULT 6
US-09-988-687-223
; Sequence 223, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/584,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-687-223

Query Match 99.0%; Score 2455.4; DB 9; Length 2908;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2465; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 ATGTGGCGCTTTGCTCGCTCTGCGGTCCGCGCGGACGACCACTGTCGAGGAGCGC 60

[illegible]

Db	1081	TTTGGGCTGACACCCAGCACCTTGGTCTCGAATGAGAACTGTGCTCAGTTTCAACAACCTT	1140
Qy	1141	CGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCGGACATCTCCCCCTCGTC	1200
Db	1141	GCAGCCACAAGATTCAACCCAGCTCAACCTCATCCACCGGACATCTCCCCCTCGTC	1200
Qy	1201	ACAGTTTTCGCTGTGAAGAGAGGCCCCACCTCAGTGTGCCATGGTTACAGGTGAA	1260
Db	1201	ACAGTTTTCGCTGTGAAGAGAGGCCCCACCTCAGTGTGCCATGGTTACAGGTGAA	1260
Qy	1261	TGCTCTCAAGTACCAGCTCGTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT	1320
Db	1261	TGCTCTCAAGTACCAGCTCGTCCAGAGGGAGTGGCAGAGGATGCCATTATTACT	1320
Qy	1321	TGCAATCCTCAGGAATTCATAGTTGAGGCGCTGCAGCTTCCAACTCCACGACAGCGTG	1380
Db	1321	TGCAATCCTCAGGAATTCATAGTTGAGGCGCTGCAGCTTCCAACTCCACGACAGTG	1380
Qy	1381	CAGGAGTACAGGAGGAGTGGCAGGACGCCCCAGCCCCAGCAGAGAAAAGTCAGTAC	1440
Db	1381	CAGGAGTACAGGAGGAGTGGCAGGACGCCCCAGCCCCAGCAGAGAAAAGTCAGTAC	1440
Qy	1441	CCAGAAATCATCTTCTTGGACAGAGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT	1500
Db	1441	CCAGAAATCATCTTCTTGGACAGAGTCTGCCATCCCGATGAAGATTCGAAATGTCAGT	1500
Qy	1501	GCCACACTGTCAACATAAGCCCGGACAGCTCTCTGCTACTGGACTGTGGTGAAGGACA	1560
Db	1501	GCCACACTGTCAACATAAGCCCGGACAGCTCTCTGCTACTGGACTGTGGTGAAGGACG	1560
Qy	1561	TTTGGCAGCTGTGCCGTATTACGGAGACCAAGTGGACAGGCTCTGGGACCCCTGGCT	1620
Db	1561	TTTGGCAGCTGTGCCGTATTACGGAGACCAAGTGGACAGGCTCTGGGACCCCTGGCT	1620
Qy	1621	GCTGTGTTGTGCCACTGCACGAGATCACACAGGGCTGCCAAGTATCTTCTG	1680
Db	1621	GCTGTGTTGTGCCACTGCACGAGATCACACAGGGCTGCCAAGTATCTTCTG	1680
Qy	1681	CAGAGAGAACGGCTTTGGCATCTTTGGGAAAGCGCTTACCCCTTCTCGTGGTGGCC	1740
Db	1681	CAGAGAGAACGGCTTTGGCATCTTTGGGAAAGCGCTTACCCCTTCTCGTGGTGGCC	1740
Qy	1741	CCCAACAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCAAGTGCAGGAGGTCTCGAC	1800
Db	1741	CCCAACAGCTCAAAGCCTGGCTCCAGCAGTACCAACACCAAGTGCAGGAGGTCTCGAC	1800
Qy	1801	CACATCAGTATGATTCCTGCCAAATGCCCTTCAAGAAAGGGGTGAGATCTCCAGTCTGCA	1860
Db	1801	CACATCAGTATGATTCCTGCCAAATGCCCTTCAAGAAAGGGGTGAGATCTCCAGTCTGCA	1860
Qy	1861	GTGGAAGAATGATCAGTTGCTGTGTTGCCAAATGTGATTTGGAAGAGTTTCAGACCTGT	1920
Db	1861	GTGGAAGAATGATCAGTTGCTGTGTTGCCAAATGTGATTTGGAAGAGTTTCAGACCTGT	1920
Qy	1921	CTGTGCGGCACCTCAAGCATCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA	1980
Db	1921	CTGTGCGGCACCTCAAGCATCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAA	1980
Qy	1981	GTGTCTATTTCCGGGACACCATGCTCCGAGGCTCTGGTCCGATGGGAAAGATGCC	2040
Db	1981	GTGTCTATTTCCGGGACACCATGCTCCGAGGCTCTGGTCCGATGGGAAAGATGCC	2040
Qy	2041	ACCTCTCTGATACATGAAGCCACCTCGAAGATGTTTGGAGAGAGAGTGCAGAAAG	2100
Db	2041	ACCTCTCTGATACATGAAGCCACCTCGAAGACGGTTTGGAGAGAGAGTGCAGAAAG	2100
Qy	2101	ACACACAGCACAACTGCCAAGCCATCAGCTGGGGATGCGGATGAAACGCGAGTTCATT	2160
Db	2101	ACACACAGCACAACTGCCAAGCCATCAGCTGGGGATGCGGATGAAACGCGAGTTCATT	2160
Qy	2161	ATGTGTAACCACTTTCAGCCAGCGCTATGCCAAGTCCCCCTCTTCAGCCCCAATTCAC	2220
Db	2161	ATGTGTAACCACTTTCAGCCAGCGCTATGCCAAGTCCCCCTCTTCAGCCCCAATTCAC	2220

```
QY 2221 GAGAAAGTGGAGTTGCCTTTTGACACATGAGGTCTGCTTTGGAGACTTTTCCACAAATG 2280
DB 2221 GAGAAAGTGGAGTTGCCTTTTGACACATGAGGTCTGCTTTGGAGACTTTTCCACAAATG 2280
QY 2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTGCGGACATCGAGAGATGGAGAG 2340
DB 2281 CCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGTGCGGACATCGAGAGATGGAGAG 2340
QY 2341 CGCAGGGAAGCGGAGTGGCGGAGGTGGCGGCGGCCCTCTCTCCAGGAGCTGGCA 2400
DB 2341 CGCAGGGAAGCGGAGTGGCGGAGGTGGCGGCGGCCCTCTCTCCAGGAGCTGGCA 2400
QY 2401 GCGGCGCTGGAGATGGGAGCCTCAGCAGAGACGGGCGGCCACACAGAGAGCCACAGGCC 2460
DB 2401 GCGGCGCTGGAGATGGGAGCCTCAGCAGAGAACGGGCGGCCACACAGAGAGCCACAGGCC 2460
QY 2461 AAGAGGTCAGAGCCAGTGA 2481
DB 2461 AAGAGGTCAGAGCCAGTGA 2481

RESULT 7
US-09-988-626-225
; Sequence 225, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtighian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988.626
; PRIORITY FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2478)
US-09-988-626-225

Query Match 98.5%; Score 2442.6; DB 9; Length 2892;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGTGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCGGACGACCATGTGCGAGGAGCGC 60
DB 1 ATGTGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCGGACGACCATGTGCGAGGAGCGC 60
QY 61 ACCATATCGCAGGACCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 ACCATATCGCAGGACCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 CGCAGCGAGAGAACCGCGGACCGTCTGCGGCTGCTCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 CGCAGCGAGAGAACCGCGGACCGTCTGCGGCTGCTCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 CAGGTGGTGGCAGCGGTAGCGGAGCTCGGCGGCGGCGGCTCTACGTCTTCTCCGAGTTC 240
DB 181 CAGGTGGTGGCAGCGGTAGCGGAGCTCGGCGGCGGCGGCTCTACGTCTTCTCCGAGTTC 240
```

```
QY 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGCAGGAGCACAAGTTA 300
DB 241 AACCGGTATCTCTTCAACTGTGGAGAGCGGTTTCAGAGACTCATGCAGGAGCACAAGTTA 300
QY 301 AAGGTTGCTCGCTCGGCAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
DB 301 AAGGTTGTTGGCTGGCAACATATTCCTGACACGAATGCACTGCTTAATGTTGGGGGC 360
QY 361 TTAAGTGAATGATTTCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 420
DB 361 TTAAGTGAATGATTTCTTACTTTAAAGAAACCGGGCTTCCAAAGTGTGACTTTCTGGA 420
QY 421 CCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTCAGAAAG 480
DB 421 CCTCCACAGCTGGAAAAATACCTCGAAGCAATCAAAATATTTCTGCTCCATTCAGAAAG 480
QY 481 ATAGAACTGGCTGTGCGGCGGCCACTCTGCGCCCGAGAAATACGAGGATGAACCATGACAGTT 540
DB 481 ATAGAACTGGCTGTGCGGCGGCCACTCTGCGCCCGAGAAATACGAGGATGAACCATGACAGTT 540
QY 541 TACAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGCAACCAACCATGGCAGAGTCCA 600
DB 541 TACCAGATCCCCATACACAGTGAAACAGAGAGGGGAAAGCAACCAACCATGGCAGAGTCCA 600
QY 601 GAAAGGCTCTTCAGCAGGCTCAGTCCAGAGCGGATCTTCAGACTCCGAGTCGAATGAAAT 660
DB 601 GAAAGGCTCTTCAGCAGGCTCAGTCCAGAGCGGATCTTCAGACTCCGAGTCGAATGAAAT 660
QY 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGTACGGGACTCTTCCCTGGTC 720
DB 661 GAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGTACGGGACTCTTCCCTGGTC 720
QY 721 GTAGCTTTCAATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCTCAAGCAAG 780
DB 721 GTAGCTTTCAATCTGTAAGCTTCACTTAAAGAGAGAAACTTCTTGGTGTCTCAAGCAAG 780
QY 781 GAGATGGGCTCCAGTGTGGACAGCTGCCATCGCTCCCATCATTCGCTGTCTCAAGAC 840
DB 781 GAGATGGGCTCCAGTGTGGACAGCTGCCATCGCTCCCATCATTCGCTGTCTCAAGAC 840
QY 841 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
DB 841 GGGAAAAGCATCACTCATGAAGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCTCCA 900
QY 901 GATCCTGGTGTCTTTTGTGGTGAATGTCCAGATGAAAGCTTCATTCACCCATC 960
DB 901 GATCCTGGTGTCTTTTGTGGTGAATGTCCAGATGAAAGCTTCATTCACCCATC 960
QY 961 TGTGAGAAATGCCACTTTTCAGAGGTACCAAGGAAAGCAGATGCCCGGCTTGGTG 1020
DB 961 TGTGAGAAATGCCACTTTTCAGAGGTACCAAGGAAAGCAGATGCCCGGCTTGGTG 1020
QY 1021 GTTCACATGGCCCGCAGCATCTGTCTTGTGACAGAGGTACCAAGAGTGGATGGAGAGG 1080
DB 1021 GTTCACATGGCCCGCAGCATCTGTCTTGTGACAGAGGTACCAAGAGTGGATGGAGAGG 1080
QY 1081 TTTGGGCTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAAACTT 1140
DB 1081 TTTGGGCTGACACCCAGCACTTGGTCTCTGAATGAGAACTGTGCTCAGTTCAAACTT 1140
QY 1141 CGGAGCCACAGATTTCAAACTGCTCAATCCACCGGACATCTTCCCGCTGCTC 1200
DB 1141 CGGAGCCACAGATTTCAAACTGCTCAATCCACCGGACATCTTCCCGCTGCTC 1200
QY 1201 ACCAGTTTCCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCCATGGTTTCAGGGTGA 1260
DB 1201 ACCAGTTTCCGCTGTAAAGAGGGGCCCCACCTCAGTGTGCCCATGGTTTCAGGGTGA 1260
QY 1261 TGCCTCTCAAGTACAGCTCCGTCAGAGGAGGTGGCAGAGGGATGCCATTATTACT 1320
DB 1261 TGCCTCTCAAGTACAGCTCCGTCAGAGGAGGTGGCAGAGGGATGCCATTATTACT 1320
```


|||||
Db 421 CCTCAGAGCTGGAATAATCTCGAGCATCAAAATATTTCTGGTCCATTGAAAGSA 480
Qy 481 ATAGAACTGGGTGTCGGGCCCACTCTCTCCCGAGAAATACGAGGATGAAACCATGACAGTT 540
Db 481 ATAGAACTGGGTGTCGGGCCCACTCTCTCCCGAGAAATACGAGGATGAAACCATGACAGTT 540
Qy 541 TACAGATCCCATACACAGTGAACAGAGGAGGGGAAAGCACCACCATGCGAGAGTCCA 600
Db 541 TACAGATCCCATACACAGTGAACAGAGGAGGGGAAAGCACCACCATGCGAGAGTCCA 600
Qy 601 GAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Db 601 GAAAGCCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCCGAATGAAAT 660
Qy 661 GAGCCACACCTTCCATCTGATGTTAGCCAGAGAGAGGGGTGAGGATCTTCTCCCTGTC 720
Db 661 GAGCCACACCTTCCATCTGATGTTAGCCAGAGAGAGGGGTGAGGATCTTCTCCCTGTC 720
Qy 721 GTAGCTTTTCATCTGTAAGCTTCACTTAAGAGAGAGAACTTCTTGGTCTCAAGCAAG 780
Db 721 GTAGCTTTTCATCTGTAAGCTTCACTTAAGAGAGAGAACTTCTTGGTCTCAAGCAAG 780
Qy 781 GAGATGGCCCTCCAGTCTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC 840
Db 781 GAGATGGCCCTCCAGTCTGGGACAGCTGCCATCGCTCCCATCATTTGCTGTCAAGGAC 840
Qy 841 GGGAAAGCATCACTCATGAAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Db 841 GGGAAAGCATCACTCATGAAAGGAGAGAGATTTTGGCTGAAGAGCTGTGTACTCTCCA 900
Qy 901 GATCTGTGTCTCTTTTGTGTGTGTAGAAATGTCAGATGAAAGCTTCAATCAACCCATC 960
Db 901 GATCTGTGTCTCTTTTGTGTGTGTAGAAATGTCAGATGAAAGCTTCAATCAACCCATC 960
Qy 961 TGTGAGATGTCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCTGCGCTTGGTG 1020
Db 961 TGTGAGATGTCACCTTTTCAGAGGTACCAAGGAAAGGAGATGCCCTGCGCTTGGTG 1020
Qy 1021 GTTCATGTGGCCCCAGCATCTGTGTGTGTGACAGAGGTACCAAGAGCTGTGTAGAGG 1080
Db 1021 GTTCATGTGGCCCCAGCATCTGTGTGTGTGACAGAGGTACCAAGAGCTGTGTAGAGG 1080
Qy 1081 TTTGGGCTGACACCCAGCATCTGTGTGTGTGATGAGACTGTGCTCAGTTTCACACCTT 1140
Db 1081 TTTGGGCTGACACCCAGCATCTGTGTGTGTGATGAGACTGTGCTCAGTTTCACACCTT 1140
Qy 1141 GCGAGCCACAAGATTCAAACCCAGCTCAACCTCAATCCACCCGACATCTTCCCTGCTC 1200
Db 1141 GCGAGCCACAAGATTCAAACCCAGCTCAACCTCAATCCACCCGACATCTTCCCTGCTC 1200
Qy 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTACGGGTGAA 1260
Db 1201 ACCAGTTTCCGCTGTAAGAGAGGGCCCCACCTCAGTGTGCCCATGGTTACGGGTGAA 1260
Qy 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTAAT 1320
Db 1261 TGCCTCTCAAGTACAGCTCCGTCAGGAGGGAGTGGCAGAGGATGCCATTATTAAT 1320
Qy 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGACAGCTTCCCACTTCAGCAGAGCGTG 1380
Db 1321 TGCAATCTGAGGAATTCATAGTTGAGGCGCTGACAGCTTCCCACTTCAGCAGAGTG 1380
Qy 1381 CAGGAGTACAGAGGAGTGCAGAGACGGCCCCAGCCCCAGCAGAGAGAAAGTCAAGTAC 1440
Db 1381 CAGGAGTACAGAGGAGTGCAGAGACGGCCCCAGCCCCAGCAGAGAGAAAGTCAAGTAC 1440
Qy 1441 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATGGAAGATTGGAATGTCAGT 1500
Db 1441 CCAGAAATCATCTTCTTGGACAGGGTCTGCCATCCCGATGGAAGATTGGAATGTCAGT 1500
Qy 1501 GCCACACTGTCAACATGAAGCCCCACACGCTCTCTGCTACTGGACTGTGGTGAGGCGACA 1560

Db 1501 GCCACACTGTCAACATAAGCCCCGACACAGCTCTCTGCTACTGGACTGTGGTGAGGCGACG 1560
Qy 1561 TTTGGGACAGCTGTCCCGTCAATTACGGAGACAGAGTGAACAGGGTCTTGGGACACCTGGCT 1620
Db 1561 TTTGGGACAGCTGTCCCGTCAATTACGGAGACAGAGTGAACAGGGTCTTGGGACACCTGGCT 1620
Qy 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCAACACAGGGCTTGGCCAAAGTATCTTGTCTG 1680
Db 1621 GCTGTGTTGTGTCCTCCACCTGACGAGATCAACACAGGGCTTGGCTAAATATCTTGTCTG 1680
Qy 1681 CAGAGAGAAAGCGGCTTGGCATCTTTGGGAAAGCGGCTTCAACCTTTGCTGGTGGTTGCC 1740
Db 1681 CAGAGAGAAAGCGGCTTGGCATCTTTGGGAAAGCGGCTTCAACCTTTGCTGGTGGTTGCC 1740
Qy 1741 CCACACAGCTCAAGGCTTGCTCCAGCATCAACACAGTGCACAGGAGGTCTGCGAC 1800
Db 1741 CCACACAGCTCAAGGCTTGCTCCAGCATCAACACAGTGCACAGGAGGTCTGCGAC 1800
Qy 1801 CACATCAGTATGATTCCTGCGCAATGCTTCAAGAGGGCTGAGATCTCCAGTCTTGCA 1860
Db 1801 CACATCAGTATGATTCCTGCGCAATGCTTCAAGAGGGCTGAGATCTCCAGTCTTGCA 1860
Qy 1861 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Db 1861 GTGAAAGATTGATCAGTTCGCTTTCGCAACATGTGATTTGGAAGAGTTTTCAGACCTGT 1920
Qy 1921 CTGTGCGGACATGCAAGCATGCGTTTGGCTGTGCGTGTGCAACCTCTGGCTGGAAA 1980
Db 1921 CTGTGCGGACATGCAAGCATGCGTTTGGCTGTGCGTGTGCAACCTCTGGCTGGAAA 1980
Qy 1981 GTGGTCTATTTCGGGGACACCATGCGCTGCGAGCTCTGGTCCGATGGGAAAGATGCC 2040
Db 1981 GTGGTCTATTTCGGGGACACCATGCGCTGCGAGCTCTGGTCCGATGGGAAAGATGCC 2040
Qy 2041 ACCCTCTGATACATGAAGCCACCTTGAAGAGAGTGGTTGGAAGAGAGTGGAAAAG 2100
Db 2041 ACCCTCTGATACATGAAGCCACCTTGAAGAGTGGTTGGAAGAGAGTGGAAAAG 2100
Qy 2101 ACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAAT 2160
Db 2101 ACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCAAT 2160
Qy 2161 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Db 2161 ATGCTGAACCACTTCAGCCAGCGTATGCCAAGTCCCTCTTTCAGCCCCCACTTCAGC 2220
Qy 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Db 2221 GAGAAAGTGGGAGTTGCTTTGACCATGAAGTCTGCTTTGGAGACTTTTCCAACAATG 2280
Qy 2281 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGCGGACATCGAGGAGATGGAGAG 2340
Db 2281 CCCAAGCTGATTTCCCACTGAAGCCCTGTTTGTGCGGCGACATCGAGGAGATGGAGAG 2340
Qy 2341 CGCAGGAGAGCGGGAGCTGCGGAGCTGCGGAGCTGCGGCGGCGCTCTCTGTCGGGAGCTGGCA 2400
Db 2341 CGCAGGAGAGCGGGAGCTGCGGAGCTGCGGAGCTGCGGCGGCGCTCTCTGTCGGGAGCTGGCA 2400
Qy 2401 GCGGCTTGGAGATGGGAGCTTCAGCAAGAGCGGGCCCCACACAGAGGAGGCCACAGGCC 2460
Db 2401 GCGGCTTGGAGATGGGAGCTTCAGCAAGAGCGGGCCCCACACAGAGGAGGCCACAGGCC 2460
Qy 2461 AAGAGGTACAGCCCAAGTGA 2481
Db 2461 AAGAGGTACAGCCCAAGTGA 2481

RESULT 9

US-09-988-626-221
; Sequence 221, Application US/09988626
; Publication No. US2003004495A1
; GENERAL INFORMATION:
; APPLICANT: Tavtligian, Sean V.

Qy	1798	CACCACATCAGTATGATTCTCGCAAAATGCCCTTCAGGAAGGGGCTGAGATCTCCAGTCCT	185
Db	1759	CACCACGTCAGTATGATTCTCGCAAAATGCCCTTCAGGAAGGGGCGAGGGTCTCCAATACT	1818
Qy	1858	GCAGTGGAAAAGATTGATTCAGTTTCGCTGTTTGCGAACATGTGATTTTGGAAAGAGTTCACAGCC	1917
Db	1819	ACATTGGAAAAGCGTGATAGACTTCGCTGTTTGGAAAACATGTGACTTAGAAGAAATTCAGACC	1878
Qy	1918	TGTCGTGTGGCGCACTCCAAAGCATCGTTTGCGCTGTGCGCTGGTGACACACCTCTGGCTGG	1977
Db	1879	TGCTGTGTACGGCACTGCAAGCATGCTTTTGGCTGTGCACATGGTACATTCATCTCTGGCTGG	1938
Qy	1978	AAAGTGTGCTATTTCGGGGNACACCATGCCCCTCGAGGCTCTGGTCCGGATCGGGAAGAT	2037
Db	1939	AAAGTCGTCTACTCGGGGATACCATGCCCTGTGAGGCTCTGGTCAGATCGGGAAAGAT	1998
Qy	2038	GCACACCTCTCTGATACATGAAGCACCCCTGGAAGATGTTTGGAAAGAGAAAGCAGTGGAA	2097
Db	1999	GCCACCTCTCTGATACATGAAGCCACTCTGGAGGATCNCCTTGGAAAGAGAAAGATAGAG	2058
Qy	2098	AAGACACAGCAACAACGTCCCAAGCCATCAGCGTGGGATGCGGATGAAACGCGAGTTC	2157
Db	2059	AGGACACACAGCACCACTCCCAAGGCTATTAATGTGGGATGCGGATGAATTCGCGAGTTC	2118
Qy	2158	ATTATGCTGAACCACTTCAGCAGCGCTATGCCAAGTCCCCCTCTTCAGCGCCCACTTC	2217
Db	2119	ATCATGCTGAACCACTTCAGTCAGCGGTACGNAAGATCCCCCTTTTCAGCCCTGACTTC	2178
Qy	2218	AGCGAGAAAGTGGGAGTTGCCTTTGACCACATCAAGAGTCTGCTTTGGAGACTTTTCCAACA	2277
Db	2179	AACGAGAAAGTTGGCATCGCCTTTGACCACATGAAGTCTGNTTTGGAGACTTCCCGACA	2238
Qy	2278	ATGCCCAAGCTGATTTCCCACTGAAGAAGCCCTGTTTGTGGCGCATCAGGAGAGATGGAG	2337
Db	2239	GTGCCCAAGCTGATTTCCCACTGAAGGCCCTGTTTGTGAGGTGCAATTTGAAGAGATGGTG	2298
Qy	2338	GAGCGCAGGGAGAAAGCGGGAGCTGCGGAGGTGCGGGCGGCCCTCTCTGTCCAGGGAGCTG	2397
Db	2299	GAACGACGGAGAAAGAGGGAGCTACGGCTGTGGCGNAGAGCCCTCTGACC---CAGCAG	2355
Qy	2398	GCAGCGGCGCTGAGGAGTGGGAGCTCTACGAGAAAGCGGGGCCCAACAGAGGAGCCACA	2456
Db	2356	GCAGACGCCACAGAGACAGAAACCCCAACAGAAAGCGGGCCCAACAGATGAACACACA	2414

RESULT 10

```

US-09-988-687-221
; Sequence 221, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtgian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

```

```
Db 979 GTGGTCCATAGCCCCAGAAATCTGTACTCATCGACAGCAGATACACAGCAGTGGATGGAG 1038
Qy 1078 AGGTTTGGGCTTGACACCCAGCAGCTTGGTCTCTGAATGAGAATCTGTCCCTCAGTTCACAAAC 1137
Db 1039 AGGTTGGGCTTGACACACAGCAGCCTGATTCTGAATGAGAAATGGCCCTCGGTCCACACAC 1098
Qy 1138 CTTGGGAGCAGAGATTTCAAAACCCAGCTCAACCTCATCCACGGAGACATCTTCCCCCTG 1197
Db 1099 CTGGCAGGCCAAGATTCAGAACCCAGCTCAGCCTCATCCACCTGACATCTTCCCCAG 1158
Qy 1198 CTCACAGTTTCGGCTGTAAGAGGAGGAGGCCACCCCTCAGTGTGCCATGTTTCAGGGT 1257
Db 1159 CTTACAGCTTCTATAGTAAGAGGAGGGTCCACCCTCAGCGTGCCAAACAGTTTCGGGT 1218
Qy 1258 GAATGCTCCTCAAGTACACAGCTTCCTCCAGGAGGAGTGGCAGAGGATGCCATTATT 1317
Db 1219 GAATGCTCCTCAAGTATTTCAGTTCGCCGCCCAAGAGAGTGGCAGAGGATACCCACACTC 1278
Qy 1318 ACTTGCAATCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGC 1377
Db 1279 GACTGCAATACTGATGAATTCATAGCTGAGGCTTGGAGCTCCCGAGTTTCCAGGAGAT 1338
Qy 1378 GTGCAGGAGTACAGAGGAGTGGCAGGAGCGGCCAGCCGCCAGCAGAGAAAGAGTCCAG 1437
Db 1339 GTGGAGGATTCGGAAGAACGTGTCAGGAAACCCAGCCGCCAGCAGAGAAAGAGCCAG 1398
Qy 1438 TACCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCATGAAGATTCGAATGTC 1497
Db 1399 TATCTCGAAATGTTCTTCTGGTACGGGTCTGCCATCCCATGGAGATCCGAATGTC 1458
Qy 1498 AGTGCCACACTTGTCAACATAGACCCCGACAGCTCTCTCTACTGACTGTGGTAGGGC 1557
Db 1459 AGTTCCACACTCTGCTCAACCTTAAGCCCTGACAAAGTCAGTCTCTGGATGTGGAGAAGC 1518
Qy 1558 ACATTGGGCAGCTGTCGGCTCATTTACGGAGACCCAGGTGGACAGGTCCTGGCCACCTG 1617
Db 1519 ACTTTTGGCAGTTGTGGCGTCAATACGGACAGCAAAATAGACGAGTCTTATGACGCTC 1578
Qy 1618 GCTGCTGTGTTGTGCCACCTGACGCGAGATCACACACGGGCTTGCACAGTATCTTG 1677
Db 1579 ACGGCTGTGTTGTGCCACCTGACGCGGACCCACACACGGGCTTGTGAATATCTG 1638
Qy 1678 CTGCAGAGAGAACGGCCTTGATCTTTGGGAAAGCCGCTTCAACCTTTGCTGTGGT 1737
Db 1639 CTGCAGAGAGAGATGCGTTGGCATCTCTGGGAAACCCCTTCAGGCCCTTGTGTGGTG 1698
Qy 1738 GCCCCCAACAGCTCAAGCCTGGCTCCAGCAGTACCAACACAGTGCAGAGGTCCTG 1797
Db 1699 GCTCCTACCCAGCTCAGGCGCTGGCTGCAGCAGATATCACAAACCTGCGCAGGAGATCTG 1758
Qy 1798 CACCACATCAGTATGATTCTTGCAAAATGCCCTTCAGGAAGGGCTGAGATCTCCAGTCT 1857
Db 1759 CACCAGCTCAGTATGATTCTGCCAAATGCCCTTCAGAAAGGGCAGAGTCTCCAACT 1818
Qy 1858 GCAGTGGAAAGATGATCAGTTCGTGTGTCGAAACATGTGATTGGAAAGATTTTCAGACC 1917
Db 1819 ACATTGGAAAGGCTGATAAGCTTGCTGTTGGAAACATGTGACTTAGAAATTTTCAGACC 1878
Qy 1918 TGTCTGTGGGCACTGCAAGCATGGTTGGCTTGGCTGTGCGCTGGTGACACCTCTGGCTGG 1977
Db 1879 TGCTTGATACGGCACTGCAAGCATGCTTTTGGCTGTGCACTGGTACATTCATCTGGCTGG 1938
Qy 1978 AAAGTGGTCTATTTCGGGGACACCATGCTCCGAGGCTCTGGTCCGATGGGAAAGAT 2037
Db 1939 AAAGTCTCTACTCGGGGATACCATGCTCTGTGAGGCTCTGGTCCAGATGGGAAAGAT 1998
Qy 2038 GCCACCTCCTGATACATGAAGCCACCTCGGAAGATGGTTTGGAAAGAGGAGCAGTGGAA 2097
Db 1999 GCCACCTCCTGATACATGAAGCCACTCTGGAGGATCNCCTTGGAGAGGAGCAGTAGAG 2058
Qy 2098 AAGACACACAGACAAAGTCCCAAGCCATCAGGTGGGATGGGATGAACGGGAGTTC 2157
```

```
Db 2059 AGGACACACAGACACACCTCCAGGCTATTATGTGGGATGCGGATGAATCGGAGTTC 2118
Qy 2158 ATTATGCTGAACCACTTTCAGCCAGCGCTATGCCAAGGTCCCTCTTTCAGCCCCAACATTC 2217
Db 2119 ATCATGCTGAACCACTTTCAGTCAGCGGTACGNAAGATCCCTTTTCAGCCCTGACTTC 2178
Qy 2218 AGCGAAGAGTGGGAGTTCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTTCCAACA 2277
Db 2179 AACGAGAAAGTGGCATCGCTTTGACCACATGAAGGTCTGNTTTGGAGACTTCCCGACA 2238
Qy 2278 ATGCCCAAGCTGATTCCCTCCACTGAAACCTGTTTGTGGCGACATCGAGGAGATGGAG 2337
Db 2239 GTGCCCAAGCTGATTCCCTCCACTGAAGCCCTGTTTTCAGGTGACATTTGAAGAGATGGTG 2298
Qy 2338 GAGCCAGGAGAAAGCGGAGCTGCGCAGGTGCGGGCGGCCCTCTGTTCAGGAGCTG 2397
Db 2299 GAACGACGGAGAGAGGAGCTACGGCTGTGTGCGAGCAGCCCTCCTGACC--CAGCAG 2355
Qy 2398 GCAGCGGCTTGAGGATGGGAGGCTCAGCAGAAAGCGGGCCACACAGAGAGCCACA 2456
Db 2356 GCAGACAGCCAGAGGACAGAGAACCCCAACAGAAAGCGGGCCACACAGATGAACACA 2414

RESULT 11
US-09-833-381-2039
; Sequence 2039, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2039
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2039
```

```
Query Match 29.6%; Score 734.8; DB 10; Length 783;
Best Local Similarity 98.3%; Pred. No. 2e-220;
Matches 772; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
```

```
Qy 91 CGGCGCGCAAGGACCCGCTGCGCACCTGCGCAGAGAGCGCGGACCGTCGGGG 150
Db 1 CGGCGCGCAAGGACCCGCTGCGCACCTGCGCAGAGAGCGCGGACCGTCGGGG 60
Qy 151 TGCTCCGCGGCCCCAAACACCGTGTACCTGCAAGTGTGGCAGCGGTAGCGCGGACTCG 210
Db 61 TGCTCCGCGGCCCCAAACACCGTGTACCTGCAAGTGTGGCAGCGGTGTA-CCGCGACTCG 119
Qy 211 GCGCGCGGCTCTAGCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAAAGGC 270
Db 120 GCGCGCGGCTCTAGCTTCTCCGAGTTCAACCGGTATCTTCAACTGTGGAGAAAGGC 179
Qy 271 GTTCAGACACTCATGCAAGGACACAGTTAAAGTTGCTCGCTCGGACAAATATTCCTG 330
Db 180 GTTCAGACACTCATGCAAGGACCAAGTTAAAGTTGCTCGCTCGGACAAATATTCCTG 239
Qy 331 ACACGAATGCAGTGTCTTAATGTTGGGGCTTAAAGTGAATGATTTCTTACTTTAAAGGAA 390
Db 240 ACACGAATGCAGTGTCTTAATGTTGGGGCTTAAAGTGAATGATTTCTTACTTTAAAGGAA 299
Qy 391 ACCGGCTTCCAAAGTGTGTACTTTCGAGCTCCCAACTGGAAAAATATCCTCGAAGCA 450
```

```
Db 300 ACCGGCTCCAAAGTGTGTAATTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCA 359
Qy 451 ATCAAAATATTTTCTGGTCCATTGAAGAAATAGAACTGGCTGTGGGCCCCACACTCTGCC 510
Db 360 ATCAAAATATTTTCTGGTCCATTGAAGAAATAGAACTGGCTGTGGGCCCCACACTCTGCC 419
Qy 511 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCCATACACAGTGAACAGAGG 570
Db 420 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCCATACACAGTGAACAGAGG 479
Qy 571 AGGGGAAGCAACCAACCATGGCAGAGTCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 630
Db 480 AGGGGAAGCAACCAACCATGGCAGAGTCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAG 539
Qy 631 CGATCTTCAGACTCCGAG-TCGAATGAATAAGAGCAACCTTCCACATGGTGTAGGCA 689
Db 540 CGATCTTCAGACTCCGAGTTCGAATGAATAAGAGCAACCTTCCACATGGTGTAGGCA 599
Qy 690 GAGAAGAGGGTCAAGGACTCTTCCCTGGTGTAGCTTTCATCTGTAAGCTTCACTTAA 749
Db 600 GAGAAGAGGGTCAAGGACTCTTCCCTGGTGTAGCTTTCATCTGTAAGCTTCACTTAA 659
Qy 750 GAGAGAAACTTCTTGGTCTCAAGCAAGAGAGATGGGCTCCAGTTGGGACAGCTGC 809
Db 660 GAGAGAAACTTCTTGGTCTCAAGCAAGAGAGATGGGCTCCCATTTGGGA-AACTGC 718
Qy 810 CATCCTCCATCATCTGCTGTCAAGACGGGAAAAGCATCATATGAAGGAAGAGA 869
Db 719 CATCCTCCATCATCTGCTGTCAAGACGACGACCAACCCACCACTAAAGGAAAAGA 778
Qy 870 GATTT 874
Db 779 GATTT 783

RESULT 12
US-09-833-381-2038
; Sequence 2038, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2038
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-2038

Query Match 19.0%; Score 470.4; DB 10; Length 536;
Best Local Similarity 99.8%; Pred. No. 2.4e-137;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 91 CGGCCGCGCAAGGACCCGCTGGCGACCTGCGCACCTGCGAGAGAGCGCGGACCGTCGGGG 150
Db 1 CGGCCGCGCAAGGACCCGCTGGCGACCTGCGCACCTGCGAGAGAGCGCGGACCGTCGGGG 60
Qy 151 TGTCTCGGGGGCCCAACACCGTGACTGTCAGGTGGTGGCAGCGGGTAGCGGGACTCG 210
Db 61 TGTCTCGGGGGCCCAACACCGTGACTGTCAGGTGGTGGCAGCGGGTAGCGGGACTCG 120
Qy 211 GCGCGCGGCTCTAGCTCTTCCAGTTCAACCGGTATCTTCAACTGTGGAGAGGC 270
Db 121 GCGCGCGGCTCTAGCTCTTCCAGTTCAACCGGTATCTTCAACTGTGGAGAGGC 180
Qy 271 GTTCAGAGACTCATGAGGAGCACAAAGTTAAAGGTTGCTCGCCTGGCAACATATTCCTG 330
.
```

```
Db 181 GTTCAGAGACTCATGAGGAGCAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTG 240
Qy 331 ACAGAAATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTACTTTAAAGGAA 390
Db 241 ACAGAAATGCACTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTACTTTAAAGGAA 300
Qy 391 ACCGGGCTTCCAAAGTGTGTACTTCTGGACCTCCAACTGGAAAAATACCTCGAAGCA 450
Db 301 ACCGGGCTTCCAAAGTGTGTACTTCTGGACCTCCAACTGGAAAAATACCTCGAAGCA 360
Qy 451 ATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 510
Db 361 ATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCC 420
Qy 511 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCATACACAGTG 562
Db 421 CCAGAAATACGAGGATGAACCAATGACAGTTTACCAAGATCCCATACACAGTG 472

RESULT 13
US-09-918-995-8996
; Sequence 8996, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8996
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(554)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8996

Query Match 17.4%; Score 432.8; DB 9; Length 554;
Best Local Similarity 91.3%; Pred. No. 1.6e-125;
Matches 502; Conservative 0; Mismatches 44; Indels 4; Gaps 4;

Qy 232 TCCAGTTCAACCGGTATCTTCAACTGTGGAGAGCGGTTCCAGAGACTCATCGAGAG 291
Db 7 TCACTGCTCGANGGGGATCGAATTTCTCTCGGAGAGCGGATCAGACATATCGAGAG 66
Qy 292 CACAAGTTAAAGGTTGCTCGCTCGCAACATATTTCTGCACCAATGCACGTGCTTAAT 351
Db 67 CACAAGTTAAAGGTTGCTCGCTCGCAACATATTTCTGCACCAATGCACGTGCTTAAT 126
Qy 352 GTTGGGGCTTAAAGTGAATGATTTCTTTAAAGGAAACCGGGCTTCCAAAGTGTA 411
Db 127 GTTGGGGCTTAAAGTGAATGATTTCTTTAAAGGAAACCGGGCTTCCAAAGTGTA 186
Qy 412 CTTTCTCGACTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCA 471
Db 187 CTTTCTCGACTCCACAACCTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCA 246
Qy 472 TTGAAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCCCAAGATACGAGGATGAAACC 531
Db 247 TTGAAAGGAATAGAACTGGCTGTGGGCCCCACACTCTGCCCAAGATACGAGGATGAAACC 306
Qy 532 ATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGG 591
Db 307 ATGACAGTTTACAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGG 366
Qy 592 CAGAGTCCAGAAAGGCTCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGACTCCGAGTCG 651
```

Db 367 CAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATCTTTCAGACTCCGAGTCG 426
QY 652 AATGAAATGAGCCACACCTTCACATGTTGTAGCCAGAGAGAGGGGTTCAGGGAACCTCT 711
Db 427 AATGAAATGAGCCACACCTTCACATGTTGTAGCCAGAGAGAGGGGTTCAGGGAACCTCT 484
QY 712 TCCTCGTGTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAGAACTTCTTGTGTGTC 771
Db 485 TCCTCGTGTAGCTTTCATCTGTAGCTTCACTTAAAGAGAGAGAACTTCTTGTGTGTC 542
QY 772 AAAGCAAGG 781
Db 543 AAAGATGG 552

RESULT 14
US-09-988-626-210
; Sequence 210, Application US/09988626
; Publication No. US2003004959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-988-626-210

Query Match 10.0%; Score 247.4; DB 9; Length 350;
Best Local Similarity 97.7%; Pred. No. 2.7e-67;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGTGGCGCTTTGCTCGTGTGCTCGGTCGCGGCGGAGCGACCATGTGCGAGGAGCG 60
Db 51 ATGTGGCGCTTTGCTCGTGTGCTCGGTCGCGGCGGAGCGACCATGTGCGAGGAGCG 110
QY 61 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 120
Db 111 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 170
QY 121 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 180
Db 171 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 230
QY 181 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 240
Db 231 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 290
QY 241 AACCGGTATCTTTCAA 257
Db 291 AACCGGTAGTCAACGA 307

Search completed: May 18, 2003, 00:05:36
Job time : 322.542 secs

RESULT 15
US-09-988-687-210
; Sequence 210, Application US/09988687
; Publication No. US20030045704A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,687
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(293)
US-09-988-687-210

Query Match 10.0%; Score 247.4; DB 9; Length 350;
Best Local Similarity 97.7%; Pred. No. 2.7e-67;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGTGGCGCTTTGCTCGTGTGCTCGGTCGCGGCGGAGCGACCATGTGCGAGGAGCG 60
Db 51 ATGTGGCGCTTTGCTCGTGTGCTCGGTCGCGGCGGAGCGACCATGTGCGAGGAGCG 110
QY 61 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 120
Db 111 ACCATATCGAGGACACCGCCCGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 170
QY 121 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 180
Db 171 CGCAGCGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 230
QY 181 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 240
Db 231 CAGGTGGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 290
QY 241 AACCGGTATCTTTCAA 257
Db 291 AACCGGTAGTCAACGA 307

GenCore version 5.1.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 15:56:54 ; Search time 3376.55 Seconds
(without alignments)
11900.029 Million cell updates/sec

Title: US-09-434-382-1
Perfect score: 2481
Sequence: 1 atgtggggcgtttctgcgt.....agaaggtcagagcccaagtga 2481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_estba:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1661	66.9	2875	11 BC023134	BC023134 Mus muscu
2	917.4	37.0	1510	11 AK004136	AK004136 Mus muscu
3	908.8	36.6	1030	14 BQ054697	BQ054697 AGENCOURT
4	876	35.3	930	14 BQ880295	BQ880295 AGENCOURT
5	872.2	35.2	918	14 BQ706437	BQ706437 AGENCOURT
6	868.8	35.0	1012	14 BQ066524	BQ066524 AGENCOURT

7	854.4	34.4	1067	14	BM927353	BM927353 AGENCOURT
8	831.2	33.5	915	14	BQ890505	BQ890505 AGENCOURT
9	828.8	33.4	1042	14	BQ057402	BQ057402 AGENCOURT
10	821.6	33.1	978	14	BQ064858	BQ064858 AGENCOURT
11	820.8	33.1	970	14	BQ071202	BQ071202 AGENCOURT
12	801.8	32.3	1009	14	BQ059218	BQ059218 AGENCOURT
13	800	32.2	983	14	BQ069865	BQ069865 AGENCOURT
14	799.4	32.2	876	14	BQ068069	BQ068069 AGENCOURT
15	799	32.2	928	14	BQ679122	BQ679122 AGENCOURT
16	793.2	32.0	881	14	BQ884266	BQ884266 AGENCOURT
17	769.6	31.0	1002	14	BM928586	BM928586 AGENCOURT
18	767	30.9	898	14	BQ923185	BQ923185 AGENCOURT
19	766.4	30.9	791	12	BQ822529	BQ822529 602725613
20	755	30.4	1065	13	B1825701	B1825701 603072685
21	751.2	30.3	930	14	BQ962958	BQ962958 AGENCOURT
22	736.2	29.7	966	12	BQ674022	BQ674022 602620010
23	731	29.5	1042	14	BM927272	BM927272 AGENCOURT
24	727.8	29.3	980	14	BQ677668	BQ677668 AGENCOURT
25	724.4	29.2	883	12	BQ746363	BQ746363 602703625
26	721.2	29.1	960	12	BF797306	BF797306 602256926
c	720.6	29.0	776	13	B1819565	B1819565 603036794
27	720.6	29.0	874	14	BQ674529	BQ674529 AGENCOURT
28	717.4	28.9	874	14	BQ674529	BQ674529 AGENCOURT
29	715.4	28.8	928	12	BQ675078	BQ675078 602621324
30	709.6	28.6	761	12	BE758820	BE758820 601590856
31	702	28.3	827	12	BQ396395	BQ396395 602459323
32	699.2	28.2	865	13	B1859833	B1859833 603386144
33	698.2	28.1	762	9	AU124671	AU124671 AU124671
c	692.4	27.9	708	14	BQ109476	BQ109476 imageqc.7
34	689.4	27.8	920	12	BQ386090	BQ386090 602455264
35	687.4	27.7	860	14	BQ642764	BQ642764 AGENCOURT
36	687.4	27.7	842	13	B1601088	B1601088 603249579
37	686.8	27.7	712	9	AU126037	AU126037 AU126037
38	686.4	27.7	1123	14	BM810990	BM810990 AGENCOURT
39	685	27.6	823	10	BE619259	BE619259 601473130
40	679	27.4	771	12	BQ766403	BQ766403 602739146
41	677.2	27.3	984	12	BE744876	BE744876 601576324
42	676.6	27.3	676	10	BE260495	BE260495 601150702
43	676	27.2	1012	12	BQ386348	BQ386348 602455550
44	674.2	27.2	1102	14	BQ056556	BQ056556 AGENCOURT
45	673.8	27.2	1102	14	BQ056556	BQ056556 AGENCOURT

ALIGNMENTS

RESULT 1
BC023134
LOCUS BC023134
DEFINITION Mus musculus, Similar to elac homolog 2 (E. coli), clone
IMAGE:4018447, mRNA.
ACCESSION BC023134
VERSION BC023134.1 GI:18605648
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2875)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/

QY 1786 CAGGAGTCTCTGACACATCATGATGATTCCTGCGCAATGCTTCAGGAAGGGCTGAG 1845
 |||||
 Db 1789 CAGGAGATTCCTGCACACCTCAGTATGATTCCTGCGCAATGCTTCAGGAAGGGCTGAG 1848
 |||||
 QY 1846 ATCTCCAGTCTCTGAGTGAAGATGATGATGCTGCTGCGACATGCTGATTTGGAA 1905
 |||||
 Db 1849 GTCTCAATACATACATGGAAGAGCTGATGAGCTTCTGTTGAACATGTCGACTTAGAA 1908
 |||||
 QY 1906 GAGTTTCAGACCTCTGCTGTGGGCACTGCAAGCATGCGCTTTGGCTGTGCGTGTGGTGCAC 1965
 |||||
 Db 1909 GAATTCAGACCTCTGCTGTGACGCACTGCAACATGCTTTGGCTGCGCACTGGTACAT 1968
 |||||
 QY 1966 ACCTCTGGCTGGAAGTGTCTATTCCGGGGACACCATGCGCTGCGAGGCTCTGTGTCGG 2025
 |||||
 Db 1969 TCATCTGGCTGGAAGTGTCTACTCGGGGATACCATGCGCTGTGAGGCTCTGTGTCAG 2028
 |||||
 QY 2026 ATGGGGAAGATGCCACCTCTCTGATACATGAGCCACCTGGAAGATGTTTGGGAAGAG 2085
 |||||
 Db 2029 ATGGGGAAGATGCCACCTCTCTGATACATGAGCCACCTGGAAGATGTTTGGGAAGAG 2088
 |||||
 QY 2086 GAAGCAGTGGAAAGACACACACACCAAGCTCCCAAGCCATCAGCTGGGGATGCGGATG 2145
 |||||
 Db 2089 GAAGCAGTAGAAGACACACACACCACTCCAGGCTATTAACTGGGGATGCGGATG 2148
 |||||
 QY 2146 AAGCGGAGTTCATATGCTGTAACCACTTCAGCCAGCGTATGCCAAGTGTCCCTCTTC 2205
 |||||
 Db 2149 AATGCGGAGTTCATATGCTGTAACCACTTCAGTCAGCGGTACGCCAAGATCCCTCTTC 2208
 |||||
 QY 2206 AGCCCAACTTCAGCAGAGAGTGGAGTTCCTTTGACCAATGAGCTCTGCTTTGA 2265
 |||||
 Db 2209 AGCCCTGACTCAACGAGAAAGTTGGCATCGCCTTTGACCAATGAGCTCTGCTTTGA 2268
 |||||
 QY 2266 GACTTTCCCAATGCTCCAGCTGATTTCCCACTGAAGCCCTGTTGCTGGCGACATC 2325
 |||||
 Db 2269 GACTTTCCCACTGCTCCAGCTGATTTCCCACTGAAGCCCTGTTGCGAGATGACAT 2328
 |||||
 QY 2326 GAGGAGATGAGAGCGCAGGAGAGAGCGGAGCTCGCGAGGTGCGGCGCCCTCTCTG 2385
 |||||
 Db 2329 GAAGAGATGTTGAAGCGCAGGAGAGAGGAGTACGGCTGTTGCGAGCAGCCCTCTG 2388
 |||||
 QY 2386 TCCAGGAGCTGCGAGCGCTGAGGATGGGAGCTTCAGCAGAGCGGCGCCACACA 2445
 |||||
 Db 2389 ACC---CAGCAGGACACAGCCAGAGGACAGAGAACCCCAAGAGCGGCGCCACACA 2445
 |||||
 QY 2446 GAGGA-----GCCACAGGCAAGAGTTCAGAGCCCAAGTGA 2481
 |||||
 Db 2446 GATGAACCAACACAGCCACACAGCAAGAGTTCAGAGCCCAAGTGA 2490
 |||||

RESULT 2
 AK004136
 LOCUS
 DEFINITION Mus musculus 18 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:110037C17:DNA segment, Chr 11, Wayne State
 University 80, expressed, full insert sequence.

ACCESSION AK004136
 VERSION AK004136.1 GI:12835201
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,
 clone_lib:RIKEN full-length enriched mouse cDNA library
 clone:110037C17.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.P., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Mazzarelli, J., Mombauts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weiss, C., Wittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5 (bases 1 to 1510)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owai, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yananaka, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-research.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LCM2040 row: p column: 17
High quality sequence stop: 647.

FEATURES
source

1..1030
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5804032"
/clone_lib="NIH MGC 99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 272 a 266 c 256 g 231 t 5 others

Query Match 36.6%; Score 908.8; DB 14; Length 1030;
Best Local Similarity 98.3%; Pred. No. 1.4e-222;
Matches 939; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 336 AATGCACGTGCTAATGTTGGGGCTTAAGTGAATGATCTTAAAGGAACCGG 395
Db 1 AATGCACGTGCTAATGTTGGGGCTTAAGTGAATGATCTTAAAGGAACCGG 60
QY 396 GCTTCCAAAGTGTACTTTCTGGACCTCCACAACCTGGAAAAATACCTCGAAGCAATCAA 455
Db 61 GCTTCCAAAGTGTACTTTCTGGACCTCCACAACCTGGAAAAATACCTCGAAGCAATCAA 120
QY 456 AATATTTTCTGTCATTGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGA 515
Db 121 AATATTTTCTGTCATTGAAAGGAATAGAACTGGCTGTGGCGGCCCACTCTGCCCCAGA 180
QY 516 ATACGAGGATGAACCATGACAGTTTACAGATCCCCATACAGTGAACAGAGGGGG 575
Db 181 ATACGAGGATGAACCATGACAGTTTACAGATCCCCATACAGTGAACAGAGGGGG 240
QY 576 AAAGCACCACCAATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 635
Db 241 AAAGCACCACCAATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 300
QY 636 TTCAGACTCCGAGTGAATGAAATAGAGCCACACCTTCCACATGGTGTAGCCAGAGAG 695
Db 301 TTCAGACTCCGAGTGAATGAAATAGAGCCACACCTTCCACATGGTGTAGCCAGAGAG 360
QY 696 AGGGGTCAAGGACTCTTCCCTGGTCTGAGCTTTCATCTGAGCTTCACTTAAAGAGAG 755
Db 361 AGGGGTCAAGGACTCTTCCCTGGTCTGAGCTTTCATCTGAGCTTCACTTAAAGAGAG 420
QY 756 AAACCTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGC 815
Db 421 AAACCTCTTGGTGTCAAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGC 480

QY 816 TCCCATCATTCGTCTCAAGGACGGAAAGAGCATCACTCATGAAGGAAGAGATTTT 875
Db 481 TCCCATCATTCGTCTCAAGGACGGAAAGAGCATCACTCATGAAGGAAGAGATTTT 540
QY 876 GGCTGAAGAGCTGTGTACTCTCTCCAGATCCTGGTGTCTGCTTTTGTGGTGAATGTCC 935
Db 541 GGCTGAAGAGCTGTGTACTCTCTCCAGATCCTGGTGTCTGCTTTTGTGGTGAATGTCC 600
QY 936 AGATGAAGAGCTTTCATCAACCCCATCTGTGAGAAATGCCACCTTTTCAGAGGTACAAGGAAA 995
Db 601 AGATGAAGAGCTTTCATCAACCCCATCTGTGAGAAATGCCACCTTTTCAGAGGTACAAGGAAA 660
QY 996 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCCCCAGCATCTGTCTGTGGACAG 1055
Db 661 GGCAGATCCCCCGTGGCTTGGTGTTCACATGCCCCCAGCATCTGTCTGTGGACAG 720
QY 1056 CAGGTACACAGCATGGATGGAGAGGTTTGGGCTGCACACCCAGCATCTTGGTCTCTGAATGA 1115
Db 721 CAGGTACACAGCATGGATGGAGAGGTTTGGGCTGCACACCCAGCATCTTGGTCTCTGAATGA 780
QY 1116 GAATGTGCTCAGTTCACAACTTCGAGGCCACAAGATTCAAAACCCAGCTCAACCTCAT 1175
Db 781 GAATGTGCTCAGTTCACAACTTCGAGGCCACAAGATTCAAAACCCAGCTCAACCTCAT 840
QY 1176 CCACCGGACATCTTCCCTCTCACCAGTTTCCGCTGTAAAGAGGGGGCCACCCCT 1235
Db 841 CCACCGGACATCTTCCCTCTCACCAGTTTCCGCTGTAAAGAGGGGGCCACCCCT 900
QY 1236 CAGTGTGCCCAT-GGTTTCAGGGTGA--TGCCTCTCTCAAGTACCAGTCCGTCCTCC 1287
Db 901 CAGTGTGCCCATGGTTTCAGGGTGAATGTCTCTCTCAAGTACCAGTCCGTCCTCC 955

RESULT 4

LOCUS BO880295 930 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8728334 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6339041
5', mRNA sequence.
ACCESSION BO880295
VERSION BO880295.1 GI:22272303
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 930)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LCM2533 row: d column: 18
High quality sequence stop: 699.

FEATURES
source

1..930
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6339041"
/clone_lib="NIH_MGC 47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT	207	a	254	c	284	g	183	t		2	others
ORIGIN											
Query Match			35.3%		Score 876;		DB 14;		Length 930;		
Best Local Similarity			98.0%;		Pred. No. 3.6e-214;				/		
Matches	896;	Conservative	,	0;	Mismatches	17;	Indels	1;	Gaps		
Qy	1534	CTGCTACTGGAGCTGTGTGAGGGCAATTTGGGCAGCTGTGCCGTCATTACGGAGACCAG	1593								
Dd	1	CTGCTACTGGAGCTGTGTGTGAGGGCAATTTGGGCAGCTGTGCCGTCATTACGGAGACCAG	60								
Qy	1594	GTGGACAGGGTCCCTGGGCACCTCGTGCTGTGTTGTGTCCTCCACCTCAGCACATCAC	1653								
Dd	61	GTGGACAGGGTCCCTGGGCACCTCGTGCTGTGTTGTGTCCTCCACCTCAGCACATCAC	120								
Qy	1654	CACACGGGCTTGGCCAAGTATCTTGCTGCAGAGAAACCGCCTTGGCATCTTTGGGAAG	1713								
Dd	121	CACACGGGCTTGGCCAAGTATCTTGCTGCAGAGAAACCGCCTTGGCATCTTTGGGAAG	180								
Qy	1714	CGCTTCACCTTTGCTGGTGTGCTGCCCCCAACAGCTCAAAGCCTGGCTCCAGCAGTAC	1773								
Dd	181	CGCTTCACCTTTGCTGGTGTGCTGCCCCCAACAGCTCAAAGCCTGGCTCCAGCAGTAC	240								
Qy	1774	CACAACAGTGCAGGAGGTCTGCACACATCAGTATGATTCCTGCCAATGCCTTCAG	1833								
Dd	241	CACAACAGTGCAGGAGGTCTGCACACATCAGTATGATTCCTGCCAATGCCTTCAG	300								
Qy	1834	GAAGGGCTGAGATCTCAGTCTCTGAGTGGAAAGATTGATCAGTTCGCTGTGTGGAAACA	1893								
Dd	301	GAAGGGCTGAGATCTCAGTCTCTGAGTGGAAAGATTGATCAGTTCGCTGTGTGGAAACA	360								
Qy	1894	TGTGATTTGGAAGAGTTTCAGACCTGCTGTGTGGCGCATCTCAAGCATGCGTTTGGCTGT	1953								
Dd	361	TGTGATTTGGAAGAGTTTCAGACCTGCTGTGTGGCGCATCTCAAGCATGCGTTTGGCTGT	420								
Qy	1954	CGCTGTGTGCACACCTCTGGCTGGAAGTGTCTATTCCGGGGACACCATGCCCTGCAG	2013								
Dd	421	CGCTGTGTGCACACCTCTGGCTGGAAGTGTCTATTCCGGGGACACCATGCCCTGCAG	480								
Qy	2014	GCTCTGTTCCGGATGGGAAGATGCCACCTCTCTGATACATGAGCCACCTCGAAGAT	2073								
Dd	481	GCTCTGTTCCGGATGGGAAGATGCCACCTCTCTGATACATGAGCCACCTCGAAGAT	540								
Qy	2074	GTTTGGGAAGAGAGCAGTGGAAAAGACACACAGCACAACTGCTCCAGGCCATCAGCGTG	2133								
Dd	541	GTTTGGGAAGAGAGCAGTGGAAAAGACACACAGCACAACTGCTCCAGGCCATCAGCGTG	600								
Qy	2134	GGATGCGGATGAACCGGAGTTCATTATGTGAACCACTTCAGCCAGCGCTATGCCAAG	2193								
Dd	601	GGATGCGGATGAACCGGAGTTCATTATGTGAACCACTTCAGCCAGCGCTATGCCAAG	660								
Qy	2194	GTCCCCCTCTTCAGCCCCAACCTTCAGCAGAAGTGGAGTTGCCCTTGACACCATGAAG	2253								
Dd	661	GTCCCCCTCTTCAGCCCCAACCTTCAGCAGAAGTGGAGTTGCCCTTGACACCATGAAG	720								
Qy	2254	GTCTGCTTTGGAGACTTTCCAAATATGCCAAGCTGATTTCCCCTCACTGAAAGCCCTGTTT	2313								
Dd	721	GTCTGCTTTGGAGACTTTCCAAATATGCCAAGCTGATTTCCCCTCACTGAAAGCCCTGTTT	780								
Qy	2314	GCTGGCGACATCGAGGATGAGAGCGCAGGAGAAAGCGGGAGCTGCGGCAAGTGGG	2373								
Dd	781	GCTGGCGACATCGAGGATGAGAGCGCAGGAGAAAGCGGGAGCTGCGGCAAGTGGG	840								
Qy	2374	CGGGCCCTCTCTGCTCCAGGAGCTGGCGGCGCTGG-AGGATGGGGAGCCTCAGCAGAA	2432								
Dd	841	CGGNCCCCCTCTCTGCTCAGGAGCTGGGCGGCTTGAAGGATGGGAGCCTCAGCAGAA	900								
Qy	,2433	CGGGCCCCACACAG	2446								

```
Db 309 AGGTTGCTCGCTGACAAACATATTCCTGACACGAAATGCATGGTCTTAATGTTGGGGC 368
|
Qy 361 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 420
|
Db 369 TTAAGTGAATGATCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTAATTTCTGGA 428
|
Qy 421 CTCCACAACTGGAAAAATACCTCGAGCAATCAAAATATTTTCTGGTCCATTTGAAGGA 480
|
Db 429 CTCCACAACTGGAAAAATACCTCGAGCAATCAAAATATTTTCTGGTCCATTTGAAGGA 488
|
Qy 481 ATAGAACTGGCTGTGGGGCCCACTCTGCCAGAAATACGAGATGAACACCATGACAGTT 540
|
Db 489 ATAGAACTGGCTGTGGGGCCCACTCTGCCAGAAATACGAGATGAACACCATGACAGTT 548
|
Qy 541 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGGCAGAGTCCA 600
|
Db 549 TACCAGATCCCATACACAGTGAACAGAGAGGGGAAAGCACCACCATGGCAGAGTCCA 608
|
Qy 601 GAAAGGCTCTCAGCAGGCTCAGTCAGACAGGATCTTCAGACTCCAGTCGAATGAAAT 660
|
Db 609 GAAAGGCTCTCAGCAGGCTCAGTCAGACAGGATCTTCAGACTCCAGTCGAATGAAAT 668
|
Qy 661 GAGCCACACTTCCACATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 720
|
Db 669 GAGCCACACTTCCACATCGTGTAGCCAGAGAGAGGGGTGAGGACTCTTCCCTGGTC 728
|
Qy 721 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGG-AAAATCTTCTGGTGTCTCAAAAGCAA 779
|
Db 729 GTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAATCTTCTGGTGTCTCAAAAGCAA 788
|
Qy 780 GGAGATGGGCTCCCGATTTGGGACAGCTGCCATCGCTCCCATCAATGCTGTGTCGAAGA 839
|
Db 789 GGAGATGGGCTCCCGATTTGGGACAGCTGCCATCGCTCCCATCAATGCTGTGTCGAAGA 848
|
Qy 840 CGGAAAAGCATCACTCATGAGGAGAGA-GAGATTTTGGCTGAAGACTGTGTAATCTTC 898
|
Db 849 CGGAAAAGCATCACTCATGAGGAGAGAGATTTTGGCTGAAGACTGTGTAATCTTC 908
|
Qy 899 CAGATCCTG 907
|
Db 909 CAAATCCTG 917
|

RESULT 6
BQ066524
LOCUS AGENCOURT_6869738 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5931173
DEFINITION 1012 bp mRNA linear EST 02-APR-2002
5', mRNA sequence.
ACCESSION BQ066524
VERSION BQ066524.1 GI:19895570
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2112 row: b column: 06
High quality sequence start: 16
High quality sequence stop: 672.
Location/Qualifiers
```

```
source
1..1012
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5931173"
/tissue_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 264 a 265 c 260 g 220 t 3 others
ORIGIN
Query Match 35.0%; Score 868.8; DB 14; Length 1012;
Best Local Similarity 96.6%; Pred. No. 2.7e-212;
Matches 950; Conservative 0; Mismatches 25; Indels 8; Gaps 6;
Qy 342 CTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTAAAGGAACCGGGCTTCC 401
|
Db 30 CTGGTCTAATGTTGGGGCTTAAGTGAATGATTTCTTAAAGGAACCGGGCTTCC 89
|
Qy 402 AAAGTGTGTACTTTCTGGACCTCCCAACTGGAATAATACCTCGAAGCAATCAAAATATT 461
|
Db 90 AAAGTGTGTACTTTCTGGACCTCCCAACTGGAATAATACCTCGAAGCAATCAAAATATT 149
|
Qy 462 TTCTGGTCCATTAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGA 521
|
Db 150 TTCTGGTCCATTAAGGAATAGAACTGGCTGTGGGGCCCACTCTGCCCCAGAAATACGA 209
|
Qy 522 GGATGAACCATGACAGTTTACAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCA 581
|
Db 210 GGATGAACCATGACAGTTTACAGATCCCCATACACAGTGAACAGAGAGGGGAAAGCA 269
|
Qy 582 CCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGA 641
|
Db 270 CCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAGA 329
|
Qy 642 CTCGAGTCCGAATGAAATAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGT 701
|
Db 330 CTCGAGTCCGAATGAAATAGCCACACCTTCCACATGGTGTAGCCAGAGAGAGGGGT 389
|
Qy 702 CAGGAGCTCTCCCTGGTGTGTAAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAACTT 761
|
Db 390 CAGGAGCTCTCCCTGGTGTGTAAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGAACTT 449
|
Qy 762 CTTGGTGTCAAAGCAAAAGGAGATGGGCTCCAGTGGGACAGCTGCCATCGTCCCAT 821
|
Db 450 CTTGGTGTCAAAGCAAAAGGAGATGGGCTCCAGTGGGACAGCTGCCATCGTCCCAT 509
|
Qy 822 CATTGCTGTGTCAGGAGCGGAAAGCATCACTCATGAAGGAAGAGAGATTTGGCTGA 881
|
Db 510 CATTGCTGTGTCAGGAGCGGAAAGCATCACTCATGAAGGAAGAGAGATTTGGCTGA 569
|
Qy 882 AGAGCTGTGTACTCTCCAGATCTCGTGTCTCTTTTGTGGTGTGTAAGTGTCCAGATGA 941
|
Db 570 AGAGCTGTGTACTCTCCAGATCTCGTGTCTCTTTTGTGGTGTGTAAGTGTCCAGATGA 629
|
Qy 942 AAGCTTTCATTAACCCATCTGTGAGAATGCCACCTTTTCAGAGGTACCAAGGAAAGGCAGA 1001
|
Db 630 AAGCTTTCATTAACCCATCTGTGAGAATGCCACCTTTTCAGAGGTACCAAGGAAAGGCAGA 689
|
Qy 1002 TGCCCCCGGT-GGCTTGTGGTGTTCATCGTGGCCCCCAGCATCTCTGTGTGTGGACAGAGT 1060
|
Db 690 TGCCCCCGGTGGGCTTGTGGTGTTCATCGTGGCCCCCAGCATCTGTGTGTGGACAGAGT 749
|
Qy 1061 ACCAGAGTGTGATGAGAGGTTTGGGCTGTGACCCAGCACTTGTGTCTGTAATGAGAACT 1120
|
```

Db 750 ACCACGATGGATGGAGAAGTTTGGGCGCTGACACCCAGCACTTGGTCTGTAATGAGAACT 809

Qy 1121 GTGCTCTAGTTCACAACTTTCGACGCCACAGATTCAAA-CCAGCTCAACTTCATCCAC 1179

Db 810 GTGCTCTAGTTCACAACTTTCGACGCCACAGATTCAAAACCCAGCTCAACTTCATCCAC 869

Qy 1180 CC-GGACATCTTCCCTCGCTCACCAGTTTCCGCT- GTAAGAGGAGGCGCCACCCCTCA 1237

Db 870 CCGGACATCTTCCCTCGCTCACCAGTTTCCGCT- GTAAGAGGAGGCGCCACCCCTCA 929

Qy 1238 GTGTCCCATGG-TTCAGGGTGAATCCCTCTCAAGTACAGCTCCGTCGCCAGGA---GG 1293

Db 930 GGGGCGCATGGTTTCAGGGTGAATCCCTCTCAGNTACAGCTCCGNTCCAAAGGA 989

Qy 1294 GAGTGGCAGAGGATGCCATTAT 1316

Db 990 AATTGGCAAAGGATGCCATTAT 1012

RESULT 7

BM927353

LOCUS

DEFINITION BM927353 1067 bp mRNA linear EST 12-MAR-2002

AGENCOURT 6704919 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5796765

5', mRNA sequence.

ACCESSION BM927353

VERSION BM927353.1 GI:19377732

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1067)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: sgaps-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LCM2022 row: a column: 22

High quality sequence stop: 650.

Location/Qualifiers

1..1067

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5796765"

/clone_lib="NIH_MGC_99"

/tissue type="lymphoma, cell line"

/lab host="DH10B (phage-resistant)"

/note="organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 277 a 280 c 266 g 237 t 7 others

ORIGIN

Query Match 34.4%; Score 854.4; DB 14; Length 1067;

Best Local Similarity 97.3%; Pred. No. 1.4e-208;

Matches 880; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

Qy 336 AATGCACTGGTCTAATGTGGGGCTTAAGTGAATGATTCTTAAAGAAACCGG 395

Db 1 AATGCACTGGTCTAATGTGGGGCTTAAGTGAATGATTCTTAAAGAAACCGG 60

Qy 396 GCTTCCAAAGTGTCTACTTTCTGCACCTCCACAACTCGAAAAATACCTCGAAGCAATCAA 455

Db 61 GCTTCCAAAGTGTCTACTTTCTGCACCTCCACAACTCGAAAAATACCTCGAAGCAATCAA 120

Qy 456 AATATTTTCTGGTCCATTTGAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCCCAGA 515

Db 121 AATATTTTCTGGTCCATTTGAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCCCAGA 180

Qy 516 ATACGAGATGAACCATGACAGTTTACCAGATCCCATACACAGTGAACAGAGAGGGG 575

Db 181 ATACGAGATGAACCATGACAGTTTACCAGATCCCATACACAGTGAACAGAGAGGGG 240

Qy 576 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 635

Db 241 AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC 300

Qy 636 TTCAGACTCCGAGTCCGAATGAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 695

Db 301 TTCAGACTCCGAGTCCGAATGAATGAGCCACACCTTCCACATGGTGTAGCCAGAGAAG 360

Qy 696 AGGGTCAGGAGCTCTTCCCTGGTCGTAGCTTTTCACTGTGAAGCTTCACTTAAAGAGAGG 755

Db 361 AGGGTCAGGAGCTCTTCCCTGGTCGTAGCTTTTCACTGTGAAGCTTCACTTAAAGAGAGG 420

Qy 756 AAATCTTCTGGTGTCAAGCAAGAGGAGATGGGCTCCAGTTCGGACAGCTGCCATCGC 815

Db 421 AAATCTTCTGGTGTCAAGCAAGAGGAGATGGGCTCCAGTTCGGACAGCTGCCATCGC 480

Qy 816 TCCCATCATTTCTGCTCAAGGACGGGAAAGCATCACTCATGAAGGAAGAGAGATTTT 875

Db 481 TCCCATCATTTCTGCTCAAGGACGGGAAAGCATCACTCATGAAGGAAGAGAGATTTT 540

Qy 876 GGCTGAAGAGCTGTGTACTCTCCAGATCCCTGGTCGTCTTTTGTGGTGTAGATGTCC 935

Db 541 GGCTGAAGAGCTGTGTACTCTCCAGATCCCTGGTCGTCTTTTGTGGTGTAGATGTCC 600

Qy 936 AGATGAAGCTTCAATTCACACCATCTGTGAAGTGCACCTTTCAGAGGTACCAAGGAAA 995

Db 601 AGATGAAGCTTCAATTCACACCATCTGTGAAGTGCACCTTTCAGAGGTACCAAGGAAA 660

Qy 996 GGCAGATCCCCCGTGGCTTTGGTGGTTTCAATGGCCCCAGCATCTGTGCTTGTGGACAG 1055

Db 661 GGCAGATCCCCCGTGGCTTTGGTGGTTTCAATGGCCCCAGCATCTGTGCTTGTGGACAG 720

Qy 1056 CAGGTACAGCAGTGGATGGAGAGTTTGGGCTGTACACCAGCATCTTGTGCTCTGAATGA 1115

Db 721 CAGGTACAGCAGTGGATGGAGAGTTTGGGCTGTACACCAGCATCTTGTGCTCTGAATGA 780

Qy 1116 GAATGTGCTCAGTTTCAACACCTTCCAGCCACAGATTCAAAACCCAGCTCAACCTCAT 1175

Db 781 GAATGTGCTCAGTTTCAACACCTTCCAGCCACAGATTCAAAACCCAGCTCAACCTCAT 840

Qy 1176 CCACCCGG---ACATCTTCCCTGCTCACCAGTTTCCGCTGTGAAGAGGAGGGCCCCAC 1232

Db 841 CCACCCGGGACATCTTCCCTGCTCAGAGTTTCCGCTGTGAAGAGGAGGGCCCCC 900

Qy 1233 CCTC 1236

Db 901 ACCC 904

RESULT 8

BO890505

LOCUS

DEFINITION BO890505 915 bp mRNA linear EST 16-AUG-2002

AGENCOURT 8071095 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6089876

5', mRNA sequence.

ACCESSION BO890505

VERSION BO890505.1 GI:22282519

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Strategy: Agencourt Bioscience Corporation
 Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LCM2330 row: f column: 21
 High quality sequence stop: 720.
 Location/Qualifiers
 1. .915
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6089876"
 /clone_lib="NIH MGC 112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."
 BASE COUNT 212 a 249 c 262 g 190 t 2 others
 ORIGIN
 Query Match 33.5%; Score 831.2; DB 14; Length 915;
 Best Local Similarity 99.5%; Pred. No. 1.2e-202;
 Matches 833; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1263 CCTCCTCAAGTACCAGCTCCGTCAGGAGGAGTGCGAGAGGATGCCATTATTACTTG 1322
 Db 1 CCTCCTCAAGTACCAGCTCCGTCAGGAGGAGTGCGAGAGGATGCCATTATTACTTG 60
 QY 1323 CAATCTGAGGAATTCATAGTTAGGCGCTGAGCTTCCCACTCCAGAGCGTGCA 1382
 Db 61 CAATCTGAGGAATTCATAGTTAGGCGCTGAGCTTCCCACTCCAGAGCGTGCA 120
 QY 1383 GGAGTACAGGAGGTCGCGAGAGCGGCCAGCCCGCAGAGAGAAAGTCACTACCC 1442
 Db 121 GGAGTACAGGAGGTCGCGAGAGCGGCCAGCCCGCAGAGAGAAAGTCACTACCC 180
 QY 1443 AGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGCAAAATGTCAGTGC 1502
 Db 181 AGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCCGATGAAGATTGCAAAATGTCAGTGC 240
 QY 1503 CACACTTGTCAATAGCCCGACACGCTCTGTCTACTGAGTGTGGTGGAGGCACATT 1562
 Db 241 CACACTTGTCAATAGCCCGACACGCTCTGTCTACTGAGTGTGGTGGAGGCACATT 300
 QY 1563 TGGGAGCTGTGCGCTCATACGGAGACAGGTGGACAGGTCCTGGGACCCCTGGCTGC 1622
 Db 301 TGGGAGCTGTGCGCTCATACGGAGACAGGTGGACAGGTCCTGGGACCCCTGGCTGC 360
 QY 1623 TGTGTTGTGTCCACCTGACGAGATCACACAGCGCTTGGCAAGTATCTTGTGCA 1682
 Db 361 TGTGTTGTGTCCACCTGACGAGATCACACAGCGCTTGGCAAGTATCTTGTGCA 420
 QY 1683 GAGAGAACCGCTTGGCATCTTTGGAAAGCGCTTCCACCTTTGCTGGTGGTGGCC 1742
 Db 421 GAGAGAACCGCTTGGCATCTTTGGAAAGCGCTTCCACCTTTGCTGGTGGTGGCC 480
 QY 1743 CAACAGCTCAAGAGCTTGGCTGAGAGTACCAACAGAGTCCAGAGAGGCTCTGCACCA 1802

Db 481 CAACAGCTCAAGAGCTTGGCTCCAGCAGTACCAACACAGTCCAGGAGGCTCTGCACCA 540
 QY 1803 CATCAGTATGATTCCTGCCAAATGCTTTCAGGAAGGGCTGAGATCTTCAGTCTGCACT 1862
 Db 541 CATCAGTATGATTCCTGCCAAATGCTTTCAGGAAGGGCTGAGATCTTCAGTCTGCACT 600
 QY 1863 GGAAGATGATCAGTCTGCTGTTGGACATGTGATTTGGAGAGGTTTCAGACCTGTCT 1922
 Db 601 GGAAGATGATCAGTCTGCTGTTGGACATGTGATTTGGAGAGGTTTCAGACCTGTCT 660
 QY 1923 GGTGCGGCACTCAAGCATGCTTGGCTGTGCGCTGTTGGACACCTCTGGCTGGAAAGT 1982
 Db 661 GGTGCGGCACTCAAGCATGCTTGGCTGTGCGCTGTTGGACACCTCTGGCTGGAAAGT 720
 QY 1983 GGTCTATTCCGGGACACCATGCTCGGAGGCTCTGGTCCGATGGGAAAGATGCCAC 2042
 Db 721 GGTCTATTCCGGGACACCATGCTCGGAGGCTCTGGTCCGATGGGAAAGATGCCAC 780
 QY 2043 CCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGGAGGAGTGGAAAA 2099
 Db 781 CCTCTGATACATGAAGCCACCTCGAAGATGTTTGAAGAGGAGGAGTGGAAAA 837
 RESULT 9
 BQ057402 1042 bp mRNA linear EST 29-MAR-2002
 LOCUS AGENCOURT_6739613 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5813014
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ057402
 VERSION BQ057402.1 GI:19816742
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1042)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lou Staudt
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Strategy: Agencourt Bioscience Corporation
 Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LCM2064 row: f column: 23
 High quality sequence stop: 663.
 Location/Qualifiers
 1. .1042
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5813014"
 /clone_lib="NIH MGC 99"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 278 a 276 c 262 g 226 t
 ORIGIN
 Query Match 33.4%; Score 828.8; DB 14; Length 1042;
 Best Local Similarity 96.9%; Pred. No. 5.4e-202;
 Matches 868; Conservative 0; Mismatches 22; Indels 6; Gaps 2;

Qy	336	AATGCACCTGGTCTAATGTTGGGGCTTAAGTGGAAATGATCTTACTTTAAAGGAACCGG	395
Db	1	AATGCACCTGGTCTAATGTTGGGGCTTAAGTGGAAATGATCTTACTTTAAAGGAACCGG	60
Qy	396	GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA	455
Db	61	GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA	120
Qy	456	AATATTTTCTGGTCAATTGAAAGGAATAGAACTGCTGTGGGCCCCCACTCTGCCCCAGA	515
Db	121	AATATTTTCTGGTCAATTGAAAGGAATAGAACTGCTGTGGGCCCCCACTCTGCCCCAGA	180
Qy	516	ATACGAGGATGAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGAGGGG	575
Db	181	ATACGAGGATGAACCATGACAGTTTACAGATCCCAATACACAGTGAACAGAGAGGGG	240
Qy	576	AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC	635
Db	241	AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC	300
Qy	636	TTCCAGACTCCGAGTCCGAATGAAATAGGCCACACCTTCCACATGGTGTAGCCAGAGA	695
Db	301	TTCCAGACTCCGAGTCCGAATGAAATAGGCCACACCTTCCACATGGTGTAGCCAGAGA	360
Qy	696	AGGGGTACGGACTCTTCCCTGGTGTAGCTTTCACTCTGTAAGCTTCACTTAAAGAGAGG	755
Db	361	AGGGGTACGGACTCTTCCCTGGTGTAGCTTTCACTCTGTAAGCTTCACTTAAAGAGAGG	420
Qy	756	AAACTTCTTGGTGCTCAAGCAAGAGATGGGCTCCAGTGGACAGCTGCCATCGC	815
Db	421	AAACTTCTTGGTGCTCAAGCAAGAGATGGGCTCCAGTGGACAGCTGCCATCGC	480
Qy	816	TCCCATCATTTGCTGTCAAGACGGGAAAGCATCACTCATGAAGGAAGAGATTTT	875
Db	481	TCCCATCATTTGCTGTCAAGACGGGAAAGCATCACTCATGAAGGAAGAGATTTT	540
Qy	876	GGCTGAAGAGCTGTACTCTCCAGATCTGGTGTCTGCTTTTGGTGGTAGATGCC	935
Db	541	GGCTGAAGAGCTGTACTCTCCAGATCTGGTGTCTGCTTTTGGTGGTAGATGCC	600
Qy	936	AGATGAAAGCTTCATTTCAACCCATCTGTGAGAAATGCCACCTTTACAGAGTACCAAGGAAA	995
Db	601	AGATGAAAGCTTCATTTCAACCCATCTGTGAGAAATGCCACCTTTACAGAGTACCAAGGAAA	660
Qy	996	GGCAGATCCCCCGTGGCTTGGTGTTTCAATGTCCTCCAGCATCTGTCTTGTGGACAG	1055
Db	661	GGCAGATCCCCCGTGGCTTGGTGTTTCAATGTCCTCCAGCATCTGTCTTGTGGACAG	720
Qy	1056	CAGGTACAGCAGTGGATGGAGAGTTTGGGCTGACACCCAGCATTTGGTCTGAATGA	1115
Db	721	CAGGTACAGCAGTGGATGGAGAGTTTGGGCTGACACCCAGCATTTGGTCTGAATGA	780
Qy	1116	GAACTGTGCTCAGTTCAACACCTTTCGACAGCCACAAAGATTCAACCCAGCTC- - -AACCT	1172
Db	781	GAACTGTGCTCAGTTCAACACCTTTCGACAGCCACAAAGATTCAACCCAGCTC- - -AACCT	840
Qy	1173	CATCCACCCGGACA- - -TCTTCCCCCTGCTCACCAGTTTCGCTGTAGAAGAGG	1225
Db	841	CATCCACCCGGACATCTTTCCCCCTGCTCACCAGTTTCGCTGTAGAAGAG	896
RESULT 10			
BQ064858			
LOCUS			
DEFINITION			
BQ064858			
5', mRNA sequence.			
ACCESSION			
BQ064858			
VERSION			
BQ064858.1			
GI:19893889			
KEYWORDS			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 978)		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
FEATURES	Contact: Robert Strausberg, Ph.D.		
	Email: rgs@bbs-rcmail.nih.gov		
	Tissue Procurement: Lou Staudt		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LCM2107 row: d column: 17		
source	High quality sequence stop: 620.		
	Location/Qualifiers		
	1..978		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5929312"		
	/clone_lib="NIH_MGC_99"		
	/tissue_type="lymphoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:		
BASE COUNT	EcoRI; cDNA made by oligo-dT priming. Directionally cloned		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	GGCAGCAG(G). Size-selected >500bp for average insert size		
	1.8kb. Library constructed by Ling Hong in the laboratory		
	of Gerald M. Rubin (University of California, Berkeley)		
	using ZAP-cDNA synthesis kit (Stratagene) and Superscript		
	II RT (Life Technologies). Note: this is a NIH_MGC		
	Library."		
	258 a 261 c 244 g 213 t 2 others		
	ORIGIN		
Query Match	33.1%; Score 821.6; DB 14; Length 978;		
	Best Local Similarity 98.5%; Pred. No. 3.7e-200;		
	Matches 849; Conservative 0; Mismatches 11; Indels 2; Gaps 2;		
Qy	336	AATGCACCTGGTCTAATGTTGGGGCTTTAAGTGGAAATGATCTTACTTTAAAGGAACCGG	395
Db	1	AATGCACCTGGTCTAATGTTGGGGCTTTAAGTGGAAATGATCTTACTTTAAAGGAACCGG	60
Qy	396	GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA	455
Db	61	GCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAA	120
Qy	456	AATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCCCAGA	515
Db	121	AATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGGGCCCCCACTCTGCCCCAGA	180
Qy	516	ATACGAGGATGAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGAGGGG	575
Db	181	ATACGAGGATGAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGAGGGG	240
Qy	576	AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC	635
Db	241	AAAGCACCACCATGGCAGAGTCCAGAAAGGCTCTCAGCAGGCTCAGTCCAGAGCGATC	300
Qy	636	TTCCAGACTCCGAGTCCGAATGAAATAGGCCACACCTTCCACATGGTGTAGCCAGAGA	695
Db	301	TTCCAGACTCCGAGTCCGAATGAAATAGGCCACACCTTCCACATGGTGTAGCCAGAGA	360
Qy	696	AGGGGTACGGACTCTTCCCTGGTGTAGCTTTCACTGTAAAGCTTCACTTTAAAGAGAGG	755
Db	361	AGGGGTACGGACTCTTCCCTGGTGTAGCTTTCACTGTAAAGCTTCACTTTAAAGAGAGG	420
Qy	756	AAACTTCTTGGTGCTCAAAAGAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC	815
Db	421	AAACTTCTTGGTGCTCAAAAGAGAGATGGGCTCCAGTGGGACAGCTGCCATCGC	480
Qy	816	TCCCATCATTTGCTGTCAAGACGGGAAAGCATCACTCATGAAGGAAGAGATTTT	875

Db 481 TCCCATCATGCTGCTCAGGACGGGAAAGCATCACTCATGAGGAGAGATTTT 540
 Qy 876 GGCTGAAGAGCTGTGTACTCTCCAGATCCTCGTGCTGCTTTTGTGGTGTAGAAATGCC 935
 Db 541 GGCTGAAGAGCTGTGTACTCTCCAGATCCTCGTGCTGCTTTTGTGGTGTAGAAATGCC 600
 Qy 936 AGATGAAGCTTCATTCACACCATCTGTGAGAAATGCCACCTTTCAGAGTACCAAGGAAA 995
 Db 601 AGATGAAGCTTCATTCACACCATCTGTGAGAAATGCCACCTTTCAGAGTACCAAGGAAA 660
 Qy 996 GGCAGATGCCCGCTGGCTGGTGTTCACATGGCCCGCCAGCATCTGCTTTGTGGACAG 1055
 Db 661 GGCAGATGCCCGCTGGCTGGTGTTCACATGGCCCGCCAGCATCTGCTTTGTGGACAG 720
 Qy 1056 CAGGTACACAGCATGGATGAGAGAGTTTGGGCTTGACACCCAGCATCTGGTCTCTGAATGA 1115
 Db 721 CAGGTACACANCAATGGATGAGAGAGTTTGGGCTTGACACCCAGCATCTGGTCTCTGAATGA 780
 Qy 1116 GAACGTGCTCAGTTCACAACTTCGCGAGCCAC-AGAATTCAACCCAGCTCAA-CTTC 1173
 Db 781 GAACGTGCTCAGTTCACAACTTCGCGAGCCACAAAGATTTCAACCCAGCTCAAACCTTC 840
 Qy 1174 ATCCACCCGGACATCTTCCCC 1195
 Db 841 ATCCACCCGGACATCTTCCCC 862

RESULT 11
 LOCUS BQ071202
 DEFINITION AGENCOURT_6853053 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927571
 5', mRNA sequence.

ACCESSION BQ071202
 VERSION BQ071202.1
 KEYWORDS GI:19900248

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 970)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LCM2102 row: 1 column: 04
 High quality sequence stop: 723.

FEATURES
 source

Location/Qualifiers
 1..970
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5927571"
 /clone_lib="NIH_MGC_47"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 215 a 269 c 298 g 186 t 2 others

ORIGIN

Query Match 33.1%; Score 820.8; DB 14; Length 970;
 Best Local Similarity 96.4%; Pred. No. 6e-200;
 Matches 894; Conservative 0; Mismatches 23; Indels 10; Gaps 5;

Qy 1534 CTGCTACTGGAGTGTGGTGAGGGACATTTGGGACGCTGTGCGCTCATTTACGGAGCCAG 1593
 Db 1 CTGCTACTGGAGTGTGGTGAGGGACATTTGGGACGCTGTGCGCTCATTTACGGAGCCAG 60
 Qy 1594 GTGGACAGGGCTCTGGGCACCTGGCTGCTGTTGTTGTCACCTGTCACCTGACCCAGATCAC 1653
 Db 61 GTGGACAGGGCTCTGGGCACCTGGCTGCTGTTGTTGTCACCTGTCACCTGACCCAGATCAC 120
 Qy 1654 CACAGCGCTTTCGCAAGATATCTTGTGTCAGAGAGAACCGCGCTTGGCATCTTTTGGGAAAG 1713
 Db 121 CACAGCGCTTTCGCAAGATATCTTGTGTCAGAGAGAACCGCGCTTGGCATCTTTTGGGAAAG 180
 Qy 1714 CCGCTTCAACCTTTCGCTGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1773
 Db 181 CCGCTTCAACCTTTCGCTGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 240
 Qy 1774 CACAAACAGTCCAGAGAGTCTCTGCACACATCATGATGATTCCTGCGCAAAATGCTTTCAG 1833
 Db 241 CACAAACAGTCCAGAGAGTCTCTGCACACATCATGATGATTCCTGCGCAAAATGCTTTCAG 300
 Qy 1834 GAAGGGCTGAGATCTCAGTCTCAGTGTGAAAGATGATCAGTTCGCTGTGCGAAACA 1893
 Db 301 GAAGGGCTGAGATCTCAGTCTCAGTGTGAAAGATGATCAGTTCGCTGTGCGAAACA 360
 Qy 1894 TGTGATTTTGGAGAGTTTTCAGACCTGCTGTCGCGGCTGCAAGCATGCGTTTGGCTGT 1953
 Db 361 TGTGATTTTGGAGAGTTTTCAGACCTGCTGTCGCGGCTGCAAGCATGCGTTTGGCTGT 420
 Qy 1954 GCGCTGTGCACACCTCTGGCTGGAAAGTGGTGTCTATTCCGGGGACACCATGCGCTGCGAG 2013
 Db 421 GCGCTGTGCACACCTCTGGCTGGAAAGTGGTGTCTATTCCGGGGACACCATGCGCTGCGAG 480
 Qy 2014 GCTCTGTCCGATGGGAAAGATGCCACCTCTCTGATACATGAAGCCACCTTGGAAAGAT 2073
 Db 481 GCTCTGTCCGATGGGAAAGATGCCACCTCTCTGATACATGAAGCCACCTTGGAAAGAT 540
 Qy 2074 GGTTTTGGAGAGGAGTGTGAAAGACACACACACACACACACACACACACACACACACAC 2133
 Db 541 GGTTTTGGAGAGGAGTGTGAAAGACACACACACACACACACACACACACACACACACAC 600
 Qy 2134 GGGATGCGGATGAACCGCGAGTTTCAATATGCTGAACCACTTCAGCCAGCGCTATGCGAAG 2193
 Db 601 GGGATGCGGATGAACCGCGAGTTTCAATATGCTGAACCACTTCAGCCAGCGCTATGCGAAG 660
 Qy 2194 GTCCCCCTCTTTCAGCCCCCAACTTCAGCGAGAAAGTGGAGTTGCGCTTTTGAACACATGAAG 2253
 Db 661 GTCCCCCTCTTTCAGCCCCCAACTTCAGCGAGAAAGTGGAGTTGCGCTTTTGAACACATGAAG 720
 Qy 2254 GTCTGCTTTGGAGACTTTTCCAAATATGCCAAGCTGATTTCCCTCCCACTGAAAGCCCTGTTT 2313
 Db 721 GTCTGCTTTGGAGACTTTTCC-ACAAATGCCCAAGCTGATTTCCCTCCCACTGAAAGCCCTGTTT 779
 Qy 2314 GCTGCGCATCTCAGAGATGGAGGAGCGCGGAGAGAGC-GGGAGCTGCGGAGCTGC- 2371
 Db 780 GCTGCGCATCTCAGAGATGGAGGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
 Qy 2372 -GGGCGGCGCTCTCTTCCA-GGGAGCTGCGAGCGCGGCGGCTTGGAGGATGGGAGCGCTC 2424
 Db 840 GGGCGGCGCTCTCTTCCA-GGGAGCTGCGAGCGCGGCGGCTTGGAGGATGGGAGCGCTC 899
 Qy 2425 CAGCAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2451
 Db 900 AAGCAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 926

RESULT 12
 BQ059218

LOCUS BQ059218 1009 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_7049753 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5815225
5', mRNA sequence.
ACCESSION BQ059218
VERSION BQ059218.1 GI:19818558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1009)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2070 row: c column: 02
High quality sequence stop: 682.
Location/Qualifiers
1..1009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5815225"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 230 a 280 c 278 g 221 t
ORIGIN
Query Match 32.3%; Score 801.8; DB 14; Length 1009;
Best Local Similarity 97.8%; Pred. No. 4.8e-195;
Matches 858; Conservative 0; Mismatches 7; Indels 12; Gaps 4;
Qy 710 CTTCCCTGCTAGCTTTTCATCTGTAGCTTCACTTAAGAGAGAACTT-CTTGGTG 768
Db 1 CTTCCCTGCTAGCTTTTCATCTGTAGCTTCACTTAAGAGAGAACTTGGTGG 60
Qy 769 CTCAAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCT 828
Db 61 CTCAAAGCAAGAGATGGGCTCCAGTTGGGACAGCTGCCATCGCTCCCATCATTTGCT 120
Qy 829 GCTGTCAAGGACGGGAAAGCATCACTCATGAAGGAAGAG-----AGATTTTGGCT 879
Db 121 GCTGTCAAGGACGGGAAAGCATCACTCATGAAGGAAGAGTTTTCAGATTTTGGCT 180
Qy 880 GAAGAGCTGTGTACTCTCCAGATCTGTGCTGCTTTTGTGGTGTAGATGTCAGAT 939
Db 181 GAAGAGCTGTGTACTCTCCAGATCTGTGCTGCTTTTGTGGTGTAGATGTCAGAT 240
Qy 940 GAAAGCTTCATTCAACCCCATCTGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGCA 999
Db 241 GAAAGCTTCATTCAACCCCATCTGTGAGATGCCACCTTTTCAGAGGTACCAAGGAAGGCA 300
Qy 1000 GATGCCCCCGTGGCTTGTGGTTACATGGCCCCCAGCATCTGTGCTTGTGGACAGCAGG 1059
Db 301 GATGCCCCCGTGGCTTGTGGTTACATGGCCCCCAGCATCTGTGCTTGTGGACAGCAGG 360

Qy 1060 TACCAGCAGTGGATGGAGAGGTTTGGGCTCACACCCAGCACTTGGTCTCTGAATGAGAAC 1119
Db 361 TACCAGCAGTGGATGGAGAGGTTTGGGCTCACACCCAGCACTTGGTCTCTGAATGAGAAC 420
Qy 1120 TGTGCTCAGTTCAAAACCTTTGCGAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCAC 1179
Db 421 TGTGCTCAGTTCAAAACCTTTGCGAGCCACAAGATTCAAAACCCAGCTCAACCTCATCCAC 480
Qy 1180 CGGACATCTTCCCTCTCTCACAGTTTCCGCTGTAGAAGGAGGGCCCACTTCAGT 1239
Db 481 CGGACATCTTCCCTCTCTCACAGTTTCCGCTGTAGAAGGAGGGCCCACTTCAGT 540
Qy 1240 GTGCCATGTTTCAGGTTGAATGCTCTCAAGTACCAGCTCCGTCGCCAGAGGAGTGG 1299
Db 541 GTGCCATGTTTCAGGTTGAATGCTCTCAAGTACCAGCTCCGTCGCCAGAGGAGTGG 600
Qy 1300 CAGAGGGATGCCATTATTACTTTCGAAATCTTGAGGAATTCATAGTTGAGCGCTGCAGCTT 1359
Db 601 CAGAGGGATGCCATTATTACTTTCGAAATCTTGAGGAATTCATAGTTGAGCGCTGCAGCTT 660
Qy 1360 CCCAATCTTCAGCAGAGCTGCAGGAGTACAGGAGGAGTGCAGGAGCGGCCAGCCCA 1419
Db 661 CCCAATCTTCAGCAGAGCTGCAGGAGTACAGGAGGAGTGCAGGAGCGGCCAGCCCA 720
Qy 1420 GCAGAGAAAGAGTCACTACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCG 1479
Db 721 GCAGAGAAAGAGTCACTACCCAGAAATCATCTTCTTGGAAACAGGGTCTGCCATCCG 780
Qy 1480 ATGAAGATTTCGAAATGTGAGTGCACACACTTGTCAACATAAGCCCC-GACAGCTCTCTGCT 1538
Db 781 ATGAAGATTTCGAAATGTGAGTGCACACACTTGTCAACATAAGCCCCGGACACGCTCTGCT 840
Qy 1539 ACTGAGCTGT-GGTGAGGGACATTTGGGAGCTGTG 1574
Db 841 ACTGAGCTGTGGTGAAGGCACATTTGGGGCCATCTG 877
RESULT 13
BQ069865
LOCUS AGENCOURT_7054570 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5922163
DEFINITION 5', mRNA sequence.
ACCESSION BQ069865
VERSION BQ069865.1 GI:19898911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 983)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2088 row: j column: 20
High quality sequence stop: 675.
Location/Qualifiers
1..983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5922163"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

EcORI; cDNA made by oligo-dT priming. Directionally
 cloned into EcORI/XhoI sites using the following 5'
 adaptor: GGCACAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library.

BASE COUNT 229 a 275 c 282 g 196 t 1 others
 ORIGIN

Query Match 32.2%; Score 800; DB 14; Length 983;
 Best Local Similarity 95.2%; Pred. No. 1.4e-194;
 Matches 869; Conservative 0; Mismatches 36; Indels 8; Gaps 4;

QY 1480 ATGAAGATTGAAATGTCAGTCCACATTTGTCAACATAAGCCCGACACGCTCTGTGCTA 1539
 Db 1 ATGAAGATTGAAATGTCAGTCCACATTTGTCAACATAAGCCCGACACGCTCTGTGCTA 60
 QY 1540 CTGGACTGTGGTGGGACACATTTGGGAGCTGTGGCGTCATTACGGAGCAGGTGGAC 1599
 Db 61 CTGGACTGTGGTGGGACACATTTGGGAGCTGTGGCGTCATTACGGAGCAGGTGGAC 120
 QY 1600 AGGGTCTGGGACACCTGGCTGCTGTTGTGTGTCCTGACCTGACGACATCACCACAG 1659
 Db 121 AGGGTCTGGGACACCTGGCTGCTGTTGTGTGTCCTGACCTGACGACATCACCACAG 180
 QY 1660 GGCCTGCCAAGTATCTTGTGCAGAGAGAACGCGCTTGGCATCTTTGGGAAGCGGCTT 1719
 Db 181 GGCCTGCCAAGTATCTTGTGCAGAGAGAACGCGCTTGGCATCTTTGGGAAGCGGCTT 240
 QY 1720 CACCCCTTGTGGTGGTCCGCCAACACAGCTCAAGCCTGCGTCAGCAGTACCACAAAC 1779
 Db 241 CACCCCTTGTGGTGGTCCGCCAACACAGCTCAAGCCTGCGTCAGCAGTACCACAAAC 300
 QY 1780 CAGTCCAGAGAGTCTGACACACATCAGTATGATTCCTGCCAATGCTTCCAGGAGG 1839
 Db 301 CAGTCCAGAGAGTCTGACACACATCAGTATGATTCCTGCCAATGCTTCCAGGAGG 360
 QY 1840 GCTGAGATCTCCAGTCTGAGTGGAAAGATTGATCAGTTCGCTGTGGGAACATGTGAT 1899
 Db 361 GCTGAGATCTCCAGTCTGAGTGGAAAGATTGATCAGTTCGCTGTGGGAACATGTGAT 420
 QY 1900 TTGGAAGAGTTTCAGACCTGTGTGGCGCACTGCAAGCATGCGTTGGCTGTGGCTG 1959
 Db 421 TTGGAAGAGTTTCAGACCTGTGTGGCGCACTGCAAGCATGCGTTGGCTGTGGCTG 480
 QY 1960 GTGCACCTCTGGTGGAAAGTGTCTATTCGGGGACACCATGCGCTGCGAGGCTCTG 2019
 Db 481 GTGCACCTCTGGTGGAAAGTGTCTATTCGGGGACACCATGCGCTGCGAGGCTCTG 540
 QY 2020 GTCCGGATGGGAAAGATGCCACCTCTGTGATACATGAAGCCACCTGGAAGATGGTTG 2079
 Db 541 GTCCGGATGGGAAAGATGCCACCTCTGTGATACATGAAGCCACCTGGAAGATGGTTG 600
 QY 2080 GAAAGGAAGCAGTGGAAAGACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATG 2139
 Db 601 GAAAGGAAGCAGTGGAAAGACACACAGCAACAGTCCCAAGCCATCAGCGTGGGGATG 660
 QY 2140 CGGATGAACGGAGTTCATTATGCTGAACCACTTCCAGCAGCGCTATGCCAAGTCCCC 2199
 Db 661 CGGATGAACGGAGTTCATTATGCTGAACCACTTCCAGCAGCGCTATGCCAAGTCCCC 720
 QY 2200 CTCCTTCAGCCCCCACTTCAGCGAG-AAAGTGGAGTTGCTTTTGACCATCAAGGTCGTG 2258
 Db 721 CTCCTTCAGCCCCCACTTCAGCGAGAAAGTGGAGTTGCTTTTGACCATCAAGGTCGTG 780
 QY 2259 -CTTTGGAGATTTCCAAATGCCCAAGCTGATTCCTCCCACTG-AAAGCCCTGTTGCT 2316
 Db 781 CTTTGGAAAACTTTCCAAATGCCCAAGCTGATTCCTCCCACTGAAAGCCCTGTTGCT 840
 QY 2317 GCGCATCTCGAGGATGGAGGA-----GCGCAGGAGAGCGGGAGCTGCGGCAAGTGC 2371

Db 841 GCGGACCTTCNAAGAGAGATGAAGACCCCGAGGGAACCGGAAGCTGCGGCAAGGGG 900
 QY 2372 GGGCGGCGCTCTCT 2384
 Db 901 CCGGGCGGCGCTCT 913

RESULT 14
 BQ688069
 LOCUS
 DEFINITION BQ688069 876 bp mRNA linear EST 15-JUL-2002
 AGENCOURT_8039759 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6210029
 5', mRNA sequence.
 ACCESSION BQ688069
 VERSION BQ688069.1 GI:21813385
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 876)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2370 row: e column: 06
 High quality sequence stop: 708.

FEATURES
 source
 1..876
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6210029"
 /clone_lib="NIH_MGC_110"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pORF7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 205 a 239 c 248 g 183 t 1 others
 ORIGIN

Query Match 32.2%; Score 799.4; DB 14; Length 876;
 Best Local Similarity 96.8%; Pred. No. 1.8e-194;
 Matches 848; Conservative 0; Mismatches 22; Indels 6; Gaps 3;

QY 1257 TGAATGCTCTCAAGTACCAGCTCCGTCACGAGGAGTGGCAGAGGATGCCATTAT 1316
 Db 1 TGAATGCTCTCAAGTACCAGCTCCGTCACGAGGAGTGGCAGAGGATGCCATTAT 60
 QY 1317 TACTTGGCAATCCTGAGGAATTCATAGTTGAGCGCTGCCAGCTTCCCACTTCCAGCAGAG 1376
 Db 61 TACTTGGCAATCCTGAGGAATTCATAGTTGAGCGCTGCCAGCTTCCCACTTCCAGCAGAG 120
 QY 1377 CGTGAGGAGTACAGGAGTGGCAGAGCGGCCACCGCCCGCAGAGAGAAAGTCA 1436
 Db 121 CGTGAGGAGTACAGGAGTGGCAGAGCGGCCACCGCCCGCAGAGAGAAAGTCA 180
 QY 1437 GTACCCAGAAATCATCTTCTTGGAAACAGGCTGCGCATCCGATCGAATTCGAAATGT 1496
 Db 181 GTACCCAGAAATCATCTTCTTGGAAACAGGCTGCGCATCCGATCGAATTCGAAATGT 240

